THE COMPARATIVE EFFECTIVENESS OF AFTER-ACTION REVIEWS IN CO-LOCATED AND DISTRIBUTED TEAM TRAINING ENVIRONMENTS

A Dissertation

by

STEVEN JARRETT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Psychology

The Comparative Effectiveness of After-Action Reviews in Co-located and Distributed

Team Training Environments

Copyright 2012 Steven Jarrett

THE COMPARATIVE EFFECTIVENESS OF AFTER-ACTION REVIEWS IN CO-LOCATED AND DISTRIBUTED TEAM TRAINING ENVIRONMENTS

A Dissertation

by

STEVEN JARRETT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Winfred Arthur, Jr.
Committee Members,	Mindy Bergman
	Michael Lindell
	Stephanie Payne
Head of Department,	Paul Wellman

August 2012

Major Subject: Psychology

ABSTRACT

The Comparative Effectiveness of After-Action Reviews in Co-located and Distributed Team Training Environments. (August 2012)

Steven Jarrett, B.A. Purdue University; M.S., Texas A&M University

Chair of Advisory Committee: Dr. Winfred Arthur, Jr.

The team-training literature provides favorable support for the after-action review (AAR)'s ability to improve cognitive, skill, and attitudinal outcomes in co-located and distributed environments. However, the *comparative* effectiveness of co-located and distributed AARs is unknown. Thus, the objective of the present study was to investigate the comparative effectiveness of co-located and distributed AARs. The present study examined the AAR's effect on performance, declarative knowledge, team-efficacy, team voice, team cohesion, and team-level reactions. Data were obtained from 492 undergraduate students (47.66% female) assigned to 123 4-person teams who participated in a team training protocol using a 3 (type of AAR review: non-AAR versus subjective AAR versus objective AAR) × 2 (geographic dispersion: co-located and distributed training environments) × 3 (sessions) repeated measures design.

The results indicate that AAR teams had significantly higher performance scores than the non-AAR teams. In addition, the AAR teams had higher perceptions of team-efficacy and higher levels of team cohesion than the non-AAR teams. With the exception of team-level reactions, there were no other significant differences between

the distributed AAR and co-located AAR conditions. Similarly, there were no significant differences across any of the outcome variables between the objective and subjective AAR conditions, indicating that the type of AAR did not impact the results of the training.

The findings of the present study highlight several practical and scientific implications that should be considered regarding AAR training. Primarily, regardless of the training environment or type of AAR, AAR training remains an effective intervention at increasing performance and attitudinal-based outcomes. In addition, the results suggest that the use of distributed AARs does not engender the proposed process losses that were hypothesized. Thus, the use of this training to reduce administrative costs may be a viable option for geographically dispersed organizations. Finally, practitioners should evaluate the extent to which increasing the amount of technology to allow for a more objective performance review, is providing the intended benefit to the trainees. The empirical research has consistently demonstrated that the use of objective review systems provides little to no benefit to the trainees. Future research is needed to determine the generalizability of these findings to other tasks, domains, team types, and levels of expertise.

DEDICATION

I would like to dedicate this dissertation and my entire graduate work to my parents who showed me the value of commitment and always supported me, regardless of the circumstances.

ACKNOWLEDGEMENTS

I would like to give a special THANK YOU to Winfred Arthur, Jr. for allowing me to test his patience. He was always there to provide me with guidance, tutelage, and encouragement as long as it did not interfere with racquetball. In addition, I would like to thank the rest of my committee who provided me with great recommendations and critiques that significantly improved this dissertation. Finally, thank you to the rest of Winfred's lab team who spent many hours in the lab collecting the data used in this study.

This research was sponsored by contracts and grants from the U.S. Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Training Research Division, Mesa, AZ awarded to Winfred Arthur, Jr.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Training is defined as "a systematic approach to learning and development to improve individual, team, or organizational effectiveness" (Goldstein & Ford, 2002, p. 22). According to Lynch and Black (1998) roughly 81% of all employers provided some means of training to their employees and this number is closer to 100% when one considers larger organizations (i.e., organizations with more than 1,000 employees). Lynch and Black (1998) also found that over a seven-year span, 57% of organizations reported expanding their formal training programs that were already in place. In addition, Noe (2010) reported that U.S. organizations spend $134 billion on training annually. Continued increases in job complexity resulting from the emergence of new and complex technologies coupled with the associated breadth and ambiguity of job demands has translated into a need for training systems to address these issues. The extant literature generally indicates that training demonstrates consistent organizational benefits such that meta-analytic research has shown that training is effective across a wide range of evaluation criteria (reaction, d = 0.60; learning, d = 0.63; behavior, d = 0.62; results, d = 0.62; Arthur, Bennett, Edens, & Bell, 2003).

Concomitant with the increase in training as a result of greater job complexity, organizations have, and continue to use more team-based systems to improve their

---

This dissertation follows the style of *Journal of Applied Psychology*.

effectiveness. Teams are defined as two or more individuals who work interdependently, have specific roles, perform specific tasks, and combine their collective efforts towards a common goal (Baker & Salas, 1996). Although dated, Devine, Clayton, Philips, Dunford, and Melner (1999) estimate that over 90% of Fortune 1000 companies use some form of work team as part of their organizational structure; this estimate is expected to be higher in the present day. Consequently, researchers and practitioners alike have devoted substantial resources to develop effective training methodologies for improving team-related knowledge and skills. One such training method is the after-action review (AAR) which has been the U.S. Army's preferred method of review following collective training—training that involves two or more individuals participating in the training session—for decades (Meliza, Bessemer, & Hiller, 1994).

The AAR, also known as the after-event review or debriefing, is an approach to training that is based on a systematic review of trainees' performance after recently completed tasks or performance episodes. In spite of its rather long history of use in both military and civilian settings, researchers have only recently begun to investigate the efficacy of the AAR as a collective training intervention and the boundary conditions under which it is most effective (Baird, Holland, & Deacon, 1999; Ellis & Davidi, 2005; Ellis, Mendel, & Aloni-Zohar, 2009; Oser, Gualtieri, Cannon-Bowers, & Salas, 1999; Ron, Lipshitz, & Popper, 2006). This research has begun to identify several design factors that may influence the effectiveness of AARs. Consequently, the present study seeks to contribute to the extant literature by investigating the comparative effectiveness of co-located and distributed AARs in enhancing team performance and processes.

Beyond their widespread use in military settings, the use of AARs in non-military organizations has steadily increased as well (Zakay, Ellis, & Shevaliski, 2004). However, as previously noted, in spite their prevalence, researchers have only recently begun to *empirically* investigate the AAR. For instance, Schurig, Jarrett, Glaze, Arthur, and Schurig (2011) were able to obtain only 15 independent samples from 7 studies that they could include in their meta-analysis of AAR effectiveness. These studies focused primarily on (a) investigating the moderating effects of specific AAR features (e.g., Alexander, Kepner, & Tregoe, 1962; Ellis & Davidi, 2005; Ellis, Ganzach, Castle, & Sekely, 2010; Zakay et al., 2004), or (b) on how to integrate technology to facilitate AAR-based training in both field and experimental settings (e.g., Prince, Salas, Brannick, & Orananu, 2005).

The present study seeks to further the extant literature in several critical ways and provide additional evaluation of the use of AARs as a collective training intervention. Specifically, the primary contribution is to provide a comparative evaluation of the effectiveness of AARs in co-located and distributed training environments. Co-located and distributed teams differ in terms of the geographic distribution of trainees, such that distributed training refers to a situation in which team members are geographically dispersed. In contrast, co-located training is one in which team members are situated in and interact in the same physical space. In addition, the present study investigates whether these effects are moderated by the objectivity of the AAR review. AAR review objectivity can be considered a continuum with no objective information on one end (i.e., subjective AAR) and only objective review information on the other end of the continuum (e.g., video-based objective AAR). For completeness, one could consider the use of diaries or memory aids to be near the midpoint on the objective review continuum. A subjective AAR in the present study refers to an AAR in which team members rely on memory to discuss their previous performance and an objective AAR in the present study refers to an AAR in which trainees have video recordings to facilitate recall, identification, and evaluation of key events that occurred during the performance episode under review.

It has been extensively documented that the effectiveness of a training intervention is moderated by its design features and characteristics (Arthur et al., 2003). Consequently, it is not only meaningful to investigate the relative effectiveness of AARs implemented in co-located and distributed training settings, but also the extent to which any potential observed effects may be moderated by the type of AAR review. Hence, empirical evidence on the efficacy of AARs under these different boundary conditions is not only recommended, but necessary (Bowers et al., 2006).

In addition, the present study also seeks to assess training effectiveness not only in terms of performance, but also in terms of additional effectiveness metrics such as declarative knowledge, team voice, team-efficacy, team cohesion, and team-level reactions, which have received limited, if any attention in the AAR literature. This provides a benefit to the AAR body of literature given that it will allow for an examination of the interrelationship between study variables. Similarly, identifying the proper pattern of relationships is necessary to ensure a consistent theoretical framework. Finally, as with any burgeoning body of literature, it is important to replicate results of past research to assess the generalizability of previous empirical findings. As such, the present study is also a constructive replication of Villado and Arthur (2012) and extends their subjective and objective AAR comparisons to distributed training environments.

Defining the AAR

At the broadest level, the AAR is an approach to training that turns a recent event into a learning opportunity by systematically reviewing the performance of a task or event of interest. The U.S. Army (1993) defines the AAR as “a professional discussion of an event, focused on performance standards, that enables soldiers to discover for themselves what happened, why it happened, and how to sustain strengths and improve on weaknesses” (p. 1). Ellis and Davidi (2005) define the AAR as “ . . . an organizational learning procedure that gives learners an opportunity to systematically analyze their behavior and to be able to evaluate the contributions of its various components to performance outcomes” (p. 857). In short, trainees systematically review their performance on a recently completed task or event.

As illustrated in Figure 1 (Villado, 2008; Villado & Arthur, 2012), in completing an AAR—with a facilitator in some instances—trainees seek answers to the following questions: What was the intended objective? What was the actual outcome? What specific actions and behaviors contributed to meeting the intended objective? What specific actions and behaviors detracted from meeting the intended objective? What is the intended future objective? What actions will increase the likelihood of meeting the intended future objective?

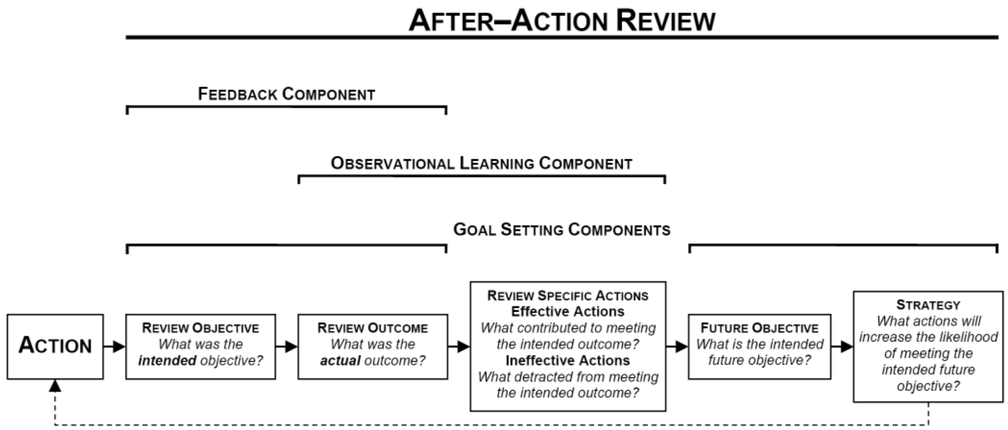

Figure 1. Primary phases of the after-action review and relevant theoretical components. In Villado, A. J. (2008). *The after-action-review training approach: An integrative framework and empirical investigation.* Unpublished dissertation, College Station, TX: Texas A&M University.

Given the long history of the AAR in military and civilian settings, one is likely to conclude that there is a strong theoretical background underlying its usage as a collective training intervention. However, in actuality, researchers' theoretical and empirical focus on the effectiveness of AARs as a training intervention has been a fairly recent occurrence. Consequently, there is little research or few theoretical models that speak to the critical elements of AARs that contribute to their effectiveness as a collective training intervention. This limited research is briefly reviewed below.

Ellis and Davidi (2005) posit that the AAR provides benefits to teams through three main functions—self-explanation, data verification, and feedback. Self-explanation is an opportunity for trainees to analyze their behavior and determine the aspects that they considered to be effective and ineffective for successful completion of the task. Data verification provides team members with an opportunity to evaluate and interpret their previous performance in different ways, allowing team members to discuss multiple perspectives of the same information. Finally, feedback provides information by which team members can re-conceptualize their understanding of the task and other important team processes (Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995).

Villado (2008) posed a similar, yet distinct, set of design features that are present in the AAR and can potentially explain the AAR's ability to improve training outcomes. Specifically, Villado (2008) posited that feedback, observational learning, and goal setting were the active mechanisms through which teams can improve performance (see Figure 1). Feedback in this situation is similar to how it was defined by Ellis and Davidi (2005). Observational learning is an opportunity for team members to review their own and teammates' behavior and determine if the current behaviors will allow for successful completion of the objectives. Finally, goal setting draws from the motivation literature wherein setting specific and challenging goals can motivate individuals towards completion of their objectives (Locke & Latham, 2002).

Over time, the basic structure of the AAR has remained relatively unchanged with the exception of the integration of various technological advances intended to facilitate the review process (e.g., video and other recording equipment to objectively document performance, and integration of recording and rating tools into simulators and simulation software). Therefore, present day AARs may differ greatly from those of the past in terms of fidelity and objectivity because of technological advances in recording, playback, and evaluation systems. However, the purpose of the AAR remains unchanged—to systematically review trainees' performance on a recent task or event in order to create a learning opportunity with the aim of improving subsequent performance.

State of the AAR Literature

In their meta-analysis, Schurig et al. (2011) found that AARs resulted in higher levels of performance (d = 1.12, k = 15) than non-AAR control conditions. When one compares the effect size for the AAR from Schurig et al. (2011) to that of other training meta-analyses, AARs fare rather favorably. Specifically, the AAR (d = 1.12, k = 15) appears to be more effective than team training in general (d = 0.87[1], k = 40; Salas et al., 2008). Additionally, the AAR was more effective than training interventions that combine goal-setting and feedback, which have yielded meta-analytic estimates ranging from d = 0.49 (Mento, Steel, & Karen, 1987) to 0.56 (Tubbs, 1986). As a caveat, it should be noted that Mento et al.'s and Tubbs' meta-analyses included both lab and field studies, whereas Schurig et al. (2011) consisted of only lab studies.

Although the extant AAR literature is limited in its breadth, there are a few notable studies that provide empirical evidence on the efficacy of various AAR design features (e.g., Alexander et al., 1962; Ellis & Davidi, 2005; Ellis et al., 2010; Ellis, Mendel, & Nir, 2006; Smith-Jentsch, Cannon-Bowers, Tannenbaum, & Salas, 2008). Alexander et al. (1962) demonstrated that structured feedback after a performance session resulted in slight improvements in performance. Although this may seem like a small gain, the control condition regressed in performance from the pre-test to post-test. So, although the AAR did not dramatically increase performance, it was capable of reducing performance declines. In addition, the effectiveness of the AAR was moderated by the extent to which the trainees were aware of the results of their actions, with the

[1] d was converted from ρ = .39.

AAR being more effective for tasks where knowledge of results was absent. That is, the AAR resulted in more performance gains (or reduced performance loss) for tasks that provided less feedback.

Moderating Effect of AAR Content

Ellis and Davidi (2005) investigated the influence of the AAR content (failure-focused or success- and failure-focused) on mental model development and the performance improvement of navigation teams. Their findings demonstrated that success- and failure-focused AARs not only resulted in more dense and complex mental models, but teams in this condition showed faster rates of performance improvement, compared to the failure-focused AAR condition. In addition, teams in the failure-focused AAR used more situational attributions to explain their performance, whereas teams that participated in success- and failure-focused AARs drew more causal links to their prior knowledge and task planning. Thus, one may posit that by drawing on both failed and successful events, teams may be more likely to ascribe the outcomes to their own actions, and not the situation.

Ellis et al. (2006) sought to extend Ellis and Davidi's (2005) work by including additional conditions (success-focused AAR and non-AAR control condition) and determining if the prior outcome of the performance episode (success or failure) interacted with the type of AAR review to influence performance. That is, did the type of AAR review moderate the relationship between a team's previous performance (i.e., prior success or prior failure) and future performance? Consistent with previous research, Ellis et al. (2006) found that AARs in general improved performance regardless of the outcome of the previous performance episode. However, the AAR *did* show differential effectiveness depending on the team's prior performance episode, such that after a failed episode all types of AAR were equally effective, but after successful performance episodes only AARs which covered failed aspects of the episode improved performance. Thus, Ellis et al. (2006) demonstrated that the AAR can provide important team-relevant information even when the team was previously successful.

In addition, in their investigation of Israeli Air Force pilots, Ron et al. (2006) found that an AAR that focused on an individual's errors served as a learning tool for all members of the team and the organization. Specifically, they found that the AAR may not only influence the individual and the team, but the entire organization's operating procedure through the identification of critical incidents discovered in the AAR. Ron et al. (2006) posit that the AAR offers an established environment to challenge the *status quo*, such that individuals and teams can learn from their own mistakes and expound on them to reduce the likelihood of similar mistakes manifesting themselves in other parts of the organization.

Moderating Effect of AAR Format

The content of the AAR is just one example of a potential moderator that has been explored in the AAR research. Another moderator that has received a substantial amount of research attention is the format of the AAR. For example, Villado et al. (2011) investigated the effectiveness of expert-led AARs compared to a non-AAR control condition in terms of performance, declarative knowledge, team-efficacy, and trainees' ability to transfer knowledge and skills. As expected, teams that participated in an expert-led AAR demonstrated higher performance and team-efficacy and were more capable of adapting their performance to a novel performance mission than non-AAR teams. However, Villado et al. (2011) did not find a difference in the declarative knowledge scores between trainees in the expert-led AAR and non-AAR conditions. The results of this study suggest that the AAR may be more effective in influencing behavioral outcomes (e.g., performance) compared to knowledge-based outcomes (e.g., declarative knowledge).

Ellis et al. (2010) investigated the effectiveness of personal AARs compared to a pre-scripted filmed AAR. In the personal AAR, the experimenter would guide the participant through a review of his or her own performance, focusing on events that were especially successful or unsuccessful. In the filmed AAR, the experimenter walked through the events similar to the personal AAR, however the performance was not based on the participant's performance; instead it was a confederate who performed the same task under the same constraints. Ellis et al.'s (2010) results indicated that individuals who participated in an AAR had significantly larger performance improvements than the individuals in the non-AAR control condition. However, contrary to expectations, the AAR format (i.e., personal versus pre-scripted) did not result in different levels of performance; the personal and pre-scripted AARs were equally effective.

Within the same context, although not a traditional AAR, Villado et al. (in press), investigated the effectiveness of an internet-based observational video rehearsal protocol in enhancing skill and knowledge retention and reacquisition after an extended period of nonuse. Specifically, a group of trainees returned to the lab once a week during an 8-

week nonuse interval to view a different task-specific refresher video. The videos were developed to refresh trainees on specific aspects of the performance task. The videos began with an example of maladaptive or ineffective performance of an event or performance sequence followed by an explanation of why it was maladaptive or ineffective, and then an example of the adaptive or effective performance of the same event or performance sequence. Similar to Ellis et al. (2010), these videos could be described as standardized filmed AARs. Villado et al.'s results indicated that the video rehearsals (i.e., recorded AARs) improved training transfer after an 8-week nonuse interval, but not retention or reacquisition.

Another potential moderator of the relationship between the AAR and performance that has received attention in the extant literature, and is related to the AAR format, is the accuracy or objectivity of the AAR (Ellis et al., 2010; Savoldelli, Naik, Park, Joo, Chow, Hamstra, 2006; Villado & Arthur, 2012). Accuracy is conceptualized here, as the extent to which ones own assessment and an external assessment of a situation are consistent (Ellis et al., 2009). When AARs are implemented in operational settings (e.g., military settings), practitioners often go to great lengths to utilize technology capable of capturing every performance event, providing an objective record of the performance episode. However, there is very limited empirical research that supports the need for this level of accuracy and objectivity. For example, the null finding between personal and filmed AARs (Ellis et al., 2010) demonstrates that even when the AAR does not present information related to the trainees' *own* performance, it remains an effective training intervention. That is, the personal AAR which one could easily postulate as being a more objective indicator of individuals' previous performance was no more effective than the pre-scripted AAR. Thus, given the costs and difficulties associated with their use, the extent to which playback systems and associated technologies are necessary is an empirical question of clear import.

The results from Ellis et al. (2009) indicate that the level of accuracy of the AAR *did* positively influence future performance, with some caveats. That is, the accuracy of participants' perceptions of performance (compared to supervisor ratings) was related to future performance only if the participant's past performance episode was successful. Hence, accuracy of performance assessment was not related to future performance for low performers. In addition, to determine the generalizability of their results, Ellis et al. (2009) implemented a lab experiment. The findings of this experiment indicated that more accurate evaluations of performance resulted in stronger positive relationships with future performance. Hence, Ellis et al. (2009) provide initial evidence demonstrating the positive relationship between feedback accuracy and performance improvement.

Villado and Arthur's (2012) study sought to investigate the relationship between the objectivity of the AAR review and team performance. Hence, Villado and Arthur compared the effectiveness of subjective and objective AARs, along with a non-AAR control group. The subjective AAR teams discussed their performance without a video playback of the previous performance episode. In contrast, the objective AAR teams discussed their previous performance with the aid of a visual playback of their previous performance episodes. So, given the nature of the feedback, it would seem that the subjective AAR teams would be more susceptible to errors of omission and commission.

Thus, subjective AARs may present a less accurate portrayal of the performance episode. Consequently, Villado and Arthur (2012) hypothesized that the objective AAR teams would demonstrate higher performance gains. However, their results indicated that although both types of AAR resulted in higher levels of performance than the control group, there were no significant differences between the two types of AARs in terms of performance. In addition, Savoldelli et al. (2006) observed a similar result in their sample of anesthesiologists who were given either oral or video-assisted oral feedback. Their results indicate that the feedback conditions had higher nontechnical performance than the control group, but there were no differences in technical performance between the two feedback conditions (i.e., oral vs video-assisted oral feedback). So, albeit small in volume, the results of the studies that have investigated the efficacy of increased accuracy and objectivity in the AAR process (Ellis et al., 2009; Villado & Arthur, 2012) have demonstrated findings contrary to the theoretical framework outlined in said studies.

The preponderance of evidence, although limited, clearly demonstrates favorable results for the AAR as an individual and team training intervention, regardless of the format or content of the AAR (Schurig et al., 2011). However, the content and format of the AAR review are only two of the several AAR design characteristics. Additional characteristics and factors that have been noted by Villado (2008) are (1) the number of trainees, (2) the training of complex versus simple tasks, (3) the provision of intrinsic versus extrinsic feedback, (4) the training of individuals versus teams, (5) whether the performance episode reviewed was a successful or failed performance episode, (6) whether training is co-located versus distributed, (7) the frequency of the AAR, (8) the spacing of the AAR, (9) the degree of structure imposed on the AAR, (10) whether the AAR is self-led or instructor- or facilitator-led, and (11) the objectivity versus subjectivity of the AAR. The present study seeks to investigate the influence of two of these factors. Specifically, it will investigate the comparative effectiveness of AARs in co-located and distributed training environments, and in addition, also investigate the extent to which the observed effects are influenced by the objectivity of the AAR review.

Geographic Distribution of Trainees

With the continuing development of faster and more powerful means of communication over large geographically dispersed locations and an increase in the globalization of organizations, there has been an increasing interest in distributed training in both the private and public sectors. Distributed training refers to training in which individuals interact simultaneously from different geographic locations (Dwyer, Oser, Salas, & Fowlkes, 2000). In the military, this may take the form of distributed mission training (DMT). This is in contrast to co-located training where team members are situated and interact in the same physical space as the other trainees and the training instructor.

One of the obvious advantages of distributed training is the ability to train team members who are geographically dispersed; thus, reducing training costs without a commensurate loss in training performance. However, this presumes that distributed training is as effective as co-located training. A critical difference between distributed and co-located training is that with distributed training, some means of technology is required to permit communication between team members. Consequently, the logistical differences between the two training environments, co-located and distributed, may impact the ability of team members to enact similar team processes such as communication, team cohesion, and team-efficacy (Daft & Lengel, 1984, 1986; Driskell, Radtke, & Salas, 2003). The type of technology is an important moderating factor of the relationship between geographic dispersion and the development of team processes because the technology dictates the potential similarity between a face-to-face and the technology-mediated interaction.

Theoretical Framework of Technology-Mediated Team Processes

Media richness theory aims to improve task performance through the effective pairing of the technology's richness and the task needs (Daft & Lengel, 1986). Richness refers to "the potential information-carrying capacity of data" (Daft & Lengel, 1984, p. 196), and technologies can be placed on a continuum based on their level of information richness. Although the technologies on Daft and Lengel's continuum do not include the advent of newer, more advanced communication technologies (e.g., video-conferencing), the theoretical structure imposed by media richness theory does not preclude their inclusion and evaluation as potential sources of information richness. Daft and Lengel (1984) use four criteria that engender the level of information richness—feedback capability, communication channels utilized, source, and language. For example, face-to-face communication is the medium with the highest information richness (when conveying qualitative information) because it allows for immediate feedback, uses both audio and visual communication channels, and is a personal source of information acquisition.

Thus, the level of similarity between face-to-face and technology-mediated communication is determined by the richness and synchronicity of the communication medium. The richness of a technology is defined in terms of the extent to which the interaction is analogous to a face-to-face interaction based on information cues and specificity (Rico & Cohen, 2005). For synchronicity, technologies range on a continuum from text-based technologies (e.g., email) with potentially low synchronicity and low richness to video-conferencing, which can provide rich interactions and high synchronicity. As per media richness theory, the type of technology chosen should match the level of task or organizational ambiguity, such that the more ambiguous the task, the higher the technology should be in information richness. That is, the more equivocal the task—a task in which information can have several interpretations—the greater the need for information rich technologies. However, empirical investigations of media richness theory have not yielded consistent positive results for identifying which technologies should be paired with which types of tasks (Daft, Lengel, & Trevino, 1987; Markus, 1994; Rice & Shook, 1990). For example, meta-analytic evidence does not support the supposition that technologies with greater information richness are more effective when performing equivocal or ambiguous tasks (Rice & Shook, 1990).

Given the lack of consistent empirical support for the media richness framework, newer theoretical frameworks for understanding the interaction between the task and the technology used in distributed environments have been proposed. For example, media synchronicity identifies several media characteristics (i.e., immediacy of feedback, symbol variety, parallelism, rehearsability, and reprocessability) which interact with the communication process (i.e., convergence and conveyance) across three group functions (i.e., production, group well-being, and member support; Dennis & Valacich, 1999). Immediacy of feedback is the extent to which the technology allows for rapid feedback regarding the communication. Symbol variety refers to the "height" of the conversation or number of different ways in which information can be conveyed using the same technology, whereas parallelism is the "width" or the number of conversations that can occur at the same time, effectively. Rehearsability is the ability of the sender to make changes and rehearse the message prior to sending it. Finally, reprocessability is the extent to which previously sent messages can be reexamined (Dennis & Valacich, 1999). Regarding the actual communication, the focus is on conveyance—the exchange of information—and convergence or the development of consistent meaning amongst receivers, which have different needs regarding the synchronicity of the communication. Thus, it would seem that conveyance is a necessary but insufficient condition for convergence and the goal of the communication has significant implications as to the technology that is necessary. When focusing on conveyance, the technology would more likely engender low synchronicity, whereas a focus on convergence would require high synchronicity. Thus, it is important to identify the needs of the task (i.e., conveyance, convergence) along with the practical implications to determine the best technology for team communication. For example, a task requiring high levels of team interdependence would likely require a technology high in synchronicity as team members would likely have a high need for conveyence to reduce role ambiguity.

Finally, Clark and Brennan (1991) developed their own set of technology characteristics that differentiate between face-to-face and technology-mediated communication. They identify six elements they deem critical to understand the potential process losses through technology-mediated interactions. Specifically, they propose that interactions vary in their degree of copresence, visibility, audibility, cotemporality, simultaneity, and sequentiality. Clark and Brennan (1991) defined copresence as the extent to which members are in the same location. Visibility is if members can see each other and audibility is if team members can hear one another. Cotemporality refers to the temporal proximity between the sent and received communication, whereas simultaneity is if members can send and receive messages simultaneously. Finally, sequentiality is the extent to which the messages are received in the correct order. For example, face-to-face team members speak and are heard in turn, but teams who communicate through email may have the order distorted if one team member misses or reads an email out of sequence.

These six communication dimensions determine the ease with which two communicators are able to understand each other’s messages and form a common representation of their messages (Clark & Brennan, 1991). The six communication dimensions place communication technologies on a continuum from face-to-face to email, with audio communication falling in the middle (Clark & Brennan, 1991). Moreover, each communication technology differs in its process and ability to coordinate a message between the sender and receiver. Similar to media synchronicity theory, Clark and Brennan (1991) posit that the ideal means of coordination in a given situation is a function of the media and the purpose of the interaction.

Using the framework proposed by Clark and Brennan (1991), Driskell et al. (2003) examined the effects of technology-mediated communication on several outcome variables (e.g., cohesiveness, communication) and concluded that in general, the lack of necessary cues (e.g., audio, visual) results in a significant reduction in the formation of these team process outcomes. They conclude that “the relative loss of contextual information in computer-mediated communication can result in greater difficulty in establishing mutual knowledge . . .” (Driskell et al., 2003, p. 317).

Specifically, within the text/video conferencing continuum is synchronous audio communication—the technology used as the main source of communication in the present study. There are certain drawbacks inherent in this technology including the absence of non-verbal cues such as body language and facial expressions. Thus, the present study posits that this loss of contextual information will manifest itself via team process loss for teams that train in distributed environments. However, this technology was used predominantly because of its generalizability to the situations (e.g., military action teams) that use similar tasks and training interventions.

Empirical Evidence of the Effectiveness of Distributed AAR

Distributed training can vastly reduce the cost of training, especially in situations where training is ongoing and continuous (e.g., military teams, firefighting teams). However, it is important that the benefits associated with reduced costs are not traded off against the effectiveness of the training intervention. Therefore, an assessment of the comparative effectiveness of AARs in co-located and distributed settings would seem to be of scientific and applied importance.

The globalization of organizations and the emergence of technologies that make distributed communication more functional and effective has resulted in a concomitant increase in distributed training research and practice. Much of this research has focused on the development of tools and evaluation metrics that can be utilized in distributed environments (Dwyer et al., 2000; Fowlkes, Lane, Salas, Franz, & Oser, 1994; Kozlowski & Bell, 2007; Roche, Kurt, & Ahrens, 2010). For example, Dwyer et al. (2000) conducted two case studies to develop a team training performance metric capable of providing a valid and reliable performance estimate in distributed training environments. Their results indicate that the targeted acceptable response to generated events or tasks (TARGET) performance methodology was able to track performance improvement over several performance episodes. Consistent with Dwyer et al. (2000), much of the research investigating the effectiveness of distributed training interventions has been conducted using quasi-experiments and case studies.

Although limited, initial evidence indicates that distributed training allows team members to effectively combine their knowledge, skills, and abilities (KSAs), without the costs of training geographically dispersed individuals in co-located settings (Dwyer et al., 2000; Salas, Oser, Cannon-Bowers, & Daskarolis-Kring, 2002; Townsend, Demarie, & Hendrickson, 1998). For example, research conducted in a distributed training environment demonstrated that event-based training was able to improve performance across several performance episodes (Dwyer et al., 2000). However, this study did not use a control group or make any comparison to co-located training, thus one cannot be confident that the improvement in performance was not a function of practice with the task per se. One domain in the distributed training literature that has been researched empirically is the AAR in terms of its effectiveness in distributed training environments (Kring, 2004; Oden, 2009).

The small body of literature investigating the effectiveness of AARs in distributed training environments indicates that AARs are an effective means of training teams in distributed settings, although their effectiveness compared to co-located AARs is less clear. For example, Kring (2004) found that team-based AARs conducted in distributed environments resulted in significant improvements over baseline performance. However, teams in the co-located condition displayed significantly higher performance than the teams in the distributed condition. Kring (2004) sought to explain the differences in performance through several communication-based processes, specifically shared mental models, team cohesion, and trust. Consonant with research demonstrating a positive relationship between shared mental models and performance (e.g., Edwards, Day, Arthur, & Bell, 2005; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Stout, Cannon-Bowers, Salas, & Milanovich, 1999), Kring (2004) posited that face-to-face teams' shared mental models would be more similar than distributed teams. He posited similar hypotheses for team cohesion and trust, such that face-to-face teams would exhibit larger amounts of team cohesion and trust than distributed teams. These hypotheses were supported, such that co-located teams demonstrated more similar shared mental models, higher team cohesion, and more trust than distributed teams; thus providing initial evidence that explains the performance differences between co-located and distributed training.

However, whereas Kring (2004) framed team cohesion and trust as potential explanatory variables for performance differences, he did not provide any direct hypotheses or statistical tests that demonstrate a relationship between these variables and performance. So, although there were differences between co-located and distributed conditions in terms of shared mental models, team cohesion, and trust, there was no direct evidence for the extent to which they explained performance differences between the conditions. Consequently, the present study seeks to provide a direct test of several process and emergent state variables' (e.g., communication, team-efficacy) relationship with geographic dispersion.

In addition to Kring (2004), Oden (2009) investigated the effect of different AAR formats in a distributed training environment. Specifically, Oden (2009) found that teams who participated in a teleconference AAR with visual feedback (analogous to the present study's objective AAR) displayed the highest performance scores, followed by the non-AAR control condition, and then the teleconference AAR condition (analogous to the present study's subjective AAR). Oden's (2009) results are at odds with what one would expect based on the preceding review of the findings in the extant literature and they may be attributed to the fact that the control condition was given time to review the training materials individually between missions (Oden, 2009). Thus, it is not unreasonable to posit that by reviewing the training materials, these participants were taking part in an individual AAR of sorts. Hence, the present study hypothesizes that the subjective (or teleconference AAR) will demonstrate higher performance than the non-AAR control condition. A different pattern of results than that of Oden (2009) is expected based on the use of a true control group that does not receive an AAR.

Oden’s (2009) findings provide mixed support for the effectiveness of AARs in distributed training environments; however the absence of a commensurate set of co-located conditions means the study does not provide any insight into the *comparative* effectiveness of the AAR in co-located and distributed training environments. Absent information to this effect, it is impossible to fully evaluate the relative efficacy of distributed and co-located AARs. That is, it is conceivable that the reduced travel and administrative costs associated with the distributed AARs may not outweigh the drawbacks associated with their potential reduced effectiveness, especially in situations where mistakes and process loss are extremely detrimental to effective performance. The previously reviewed literature speaks to the efficacy of the AAR as a training intervention, but it also demonstrates how the process loss of distributed teams can lead to performance decrements. Thus, it is posited that there will be a significant main effect for geographic dispersion and AAR conditions, such that:

H1(a): Teams in the co-located and distributed AAR condition will have higher performance scores than teams in the co-located and distributed non-AAR condition.

H1(b): Teams in the co-located AAR conditions will have higher performance scores than teams in the distributed AAR conditions.

Objective versus Subjective AARs

A large proportion of the AAR literature has been concerned with developing and assessing the ease of incorporating and using objective review systems during AARs. AARs rely largely on the identification and evaluation of behaviors and key events that occurred during the performance episode under review. Technology is often integrated into the review process to facilitate the identification of critical incidents via a more objective review process. The objectivity of a review process can be considered a continuum and the present study investigates the two extremes of the continuum (i.e., subjective and objective AAR reviews). “Subjective” AARs rely on the memory and the ability of trainees (and the instructor) to recall the behaviors and events under review. Thus, the subjective AAR has little to no objective performance information. “Objective” AARs on the other hand, use recordings (e.g., video, audio, written communication, flight data recorders) to facilitate the recall, identification, and evaluation of behaviors or key events that occurred during the performance episode under review.

As previously noted, an additional objective of the present study is to investigate the effect of objectivity on the comparative effectiveness of co-located and distributed AARs. Said effect has important implications for the design of AAR training and the environments in which they are used. Although some environments may lend themselves to objective reviews, other environments may be less able to facilitate objectivity without the aid of cost-prohibitive or intrusive tools. The extent to which objectivity enhances the effectiveness of AARs would thus yield useful information regarding the utility of implementing such an approach in an environment where objective information is difficult, costly, or impossible to obtain.

It appears that practice has outpaced science in that empirical investigations examining the effect of objective review systems is lacking. Although recent work has investigated the efficacy of having trainees watch other trainees participate in an AAR (Ellis et al., 2010) or even standardized filmed AARs (Villado et al., in press), Villado and Arthur (2012) and Oden (2009) are the only two studies that have directly tested the relationship between AAR objectivity and performance improvement. However, Villado and Arthur (2012) was conducted with only co-located trainees and Oden (2009) used only distributed trainees. Thus, there is no investigation of the *comparative* effectiveness of co-located and distributed AARs.

In spite of the limited research investigating the effect of objectivity in the AAR review process, the performance appraisal and assessment center literatures provide some insights into the potential effects of objectivity of AARs, specifically in terms of the effect of objectivity on the comprehensiveness and accuracy of evaluations and assessments (DeNisi, Robbins, & Cafferty, 1989; Sturman, Cheramie, & Cashen, 2005).

Performance appraisals conducted with and without memory aids (e.g., diaries and notes) often demonstrate similar levels of rating accuracy in terms of assessing the overall performance of a target (Middendorf & Macan, 2002; Ryan et al., 1995; Sanchez & De La Torre, 1996; Woehr & Feldman, 1993). Similar results have been reported in the assessment center literature (e.g., Ryan et al., 1995). Assessors who make their ratings during the exercise are just as accurate as assessors who view videotaped recordings of the exercises (Ryan et al., 1995). However, research has also demonstrated that performance ratings made from memory have less recall accuracy (Middendorf & Macan, 2002; Ryan et al., 1995; Sanchez & De La Torre, 1996; Woehr & Feldman, 1993). For example, DeNisi et al. (1989) found that raters who relied on their memory when evaluating a target recalled fewer incidents and made more recall errors than those who were allowed to supplement their recall with a diary-like aid. This finding is noteworthy given the relatively short time interval (i.e., several minutes) between the observation of performance and the rating session. Similar research has demonstrated that errors in recall are more pronounced as the time interval between the performance episodes and the rating increases (DeNisi et al., 1989; Murphy & Balzer, 1986; Williams, DeNisi, Meglino, & Cafferty, 1986). Like DeNisi et al. (1989), Ryan et al. (1995) also found that providing assessors with access to videotaped recordings of the exercises resulted in a greater quality of recorded observations.

Taken together, it would appear that raters are able to form general evaluations while observing behavior, and are able to provide subsequent ratings based on those general evaluations (Murphy & Balzer, 1986; Woehr & Feldman, 1993). When raters do not, or are not able to form internal evaluations prior to providing a formal rating, raters base their ratings on memory (Murphy & Balzer, 1986; Woehr & Feldman, 1993). As the time interval between performance and the rating of said performance increases, ratings are less influenced by the details of the performance and more influenced by the general impression formed about the performance (Murphy & Balzer, 1986). So, recall accuracy may not be as critical for accurate appraisals as one might expect. If raters are able to form accurate internal evaluations prior to the appraisal and are able to remember those internal evaluations, they simply need to draw on those appraisals when rating the target (Murphy & Balzer, 1986; Woehr & Feldman, 1993).

Although it may be feasible for experienced supervisors to accurately recall events or generate memory aids during their workday to be used for subsequent performance evaluations, it is less likely that trainees would be able to do the same *during* training. Trainees may be less able to identify and encode or physically record critical incidents and less able to form evaluations about their performance than more experienced individuals because of the cognitive demands of learning new tasks and simultaneously attending to and evaluating their performance. It is even less likely that trainees would be able to simultaneously record and evaluate the performance of other trainees during a task given the difficulty of noting and evaluating their own performance.

Precise recall of critical incidents is needed for effective performance feedback (Murphy, 1991). Errors in recall may lead trainees to omit behaviors or key events that affected performance. Errors may also lead trainees to include irrelevant behaviors or events in the review. Recall errors that result in review deficiencies or contamination may impede the effectiveness of the AAR. Because trainees are less able to generate memory aids for their behavior and the behavior of others, objective review methods, other than self-generated diaries, may enhance the training effectiveness of AAR-based training. Thus, given the previous discussion of the benefits of objectivity in feedback, it is proposed that there will be a significant main effect for the type of AAR review, such that:

H1(c): Teams in the co-located and distributed objective AAR conditions will display higher performance than teams in the co-located and distributed subjective AAR conditions.

In summary, although there is some initial evidence that speaks to the efficacy of different types of AAR reviews in various training environments (e.g., Kring, 2004; Oden, 2009; Villado & Arthur, 2012), further work is needed to determine the generalizability of these findings, as well as the comparative effectiveness of the different types of AAR reviews in co-located and distributed training environments. As such, the objective of the present study is to provide a comparative evaluation of the effectiveness of different types of AAR reviews (i.e., subjective and objective AARs) in co-located and distributed training environments in terms of various training evaluation outcomes, specifically performance, declarative knowledge, team voice, team cohesion, team-efficacy, and team-level reactions to the training.

The Effect of AARs on Process and Outcome Variables

The objective of any team training intervention is to improve trainees' standing on the outcome variable of interest. In the context of team training, effectiveness can be conceptualized in terms of cognitive, process, performance, and affective/attitudinal outcomes (Salas et al., 2008). Researchers have a vested interest in identifying team training interventions that can improve team performance, but a focus on only performance is deficient when attempting to truly understand the effectiveness of a particular team training intervention. Therefore, in the absence of empirical research demonstrating the efficacy of training interventions in distributed environments across a wide range of outcome variables, their effectiveness in these situations is unknown. So, whereas the apparent advantages of distributed training in the military and other organizations is often highlighted, it must be emphasized that there is little research that provides a guiding framework for the development of sound methodologies that can appropriately leverage the capabilities of distributed training protocols (Oser et al., 1999). Thus, to provide a more comprehensive investigation of the comparative effectiveness of the AAR in co-located and distributed training environments, the present study examines the effectiveness of the AAR across several training effectiveness outcomes. That is, in addition to performance, researchers are often interested in other outcomes such as knowledge acquisition, team-efficacy, team voice, team cohesion, and team-level reactions.

Examining the efficacy of AARs across a variety of team effectiveness outcomes is of particular import when using the AAR in a novel training environment. The following sections of the chapter present a detailed discussion of the theoretical and conceptual framework for each of the training effectiveness variables followed by the specific hypotheses for each. Due to the limited amount of empirical research on distributed training—in the context of action teams performing interdependent tasks—the theories used here will draw from the virtual team literature as warranted.

Teamwork vs Taskwork Processes

The team performance literature distinguishes between two dimensions of team behavior, teamwork skills and taskwork or technical skills. Specifically, teamwork skills are considered global KSAs necessary for individuals to perform interdependently towards a common team goal. In addition, teamwork skills are considered behaviors required for cooperative functioning. Teamwork skills are distinct from taskwork or technical skills, in that taskwork skills are task/job specific (Kozlowski & Bell, 2003). Previous research in this field has identified the predictive ability of taskwork skills using tools such as declarative knowledge tests (Banks & Millwood, 2007) and land navigation skills for military personnel (Goodwin, 1999) to predict future performance.

The relationship between team performance and teamwork skills including interpersonal relations, communication, and decision-making has also been previously established in the literature (Cannon-Bowers & Salas, 1997). Independent of the taskwork skills necessary to perform in a specific domain, developing teamwork skills may positively influence and be a necessary condition for superior team performance (Ellis, Bell, Ployhart, Hollenbeck, & Ilgen, 2005; Salas, Bowers, & Rhodenizer, 1998). For example, Ellis et al. (2005) found that teamwork specific training was able to improve both cognitive and skill-based outcomes in a command-and-control task. In addition, Rapp and Mathieu (2007) found that teams that participated in taskwork and teamwork training sessions performed significantly better in a market simulation task than teams that were only given information on the technical knowledge necessary to perform the task.

These findings indicate that to maximize team performance, members must not only understand the task domain, they must also know how to effectively interact as a team. Although teamwork skills are considered an important facet of team performance, it would seem teamwork is a necessary but not sufficient condition for high performing teams. This is not to underplay the importance of teamwork skills, but instead to highlight the significance of taskwork skills for effective performance in a task/job domain. The present study investigates the importance of both teamwork (e.g., team voice and cohesion) and taskwork (performance) team processes. However, the AAR is being conducted in the present study is specifically designed to improve taskwork and has less of a focus on teamwork.

Team-Efficacy

Team-efficacy refers to a team's "shared belief in its conjoint capabilities to organize and execute the courses of action required to produce given levels of attainments" (Bandura, 1997, p. 477). The relationship between team-efficacy and team performance has been well-documented (e.g., Arthur, Edwards, & Bell, 2007; Gully, Incalcaterra, Joshi, & Beaubien, 2002; Lindsley, Brass, & Thomas, 1995; Porter, 2005). The primary antecedents of self-efficacy (and similarly team-efficacy) consist of enactive mastery (experience), vicarious experience, verbal persuasion, and physiological states (Bandura, 1997). Given these antecedents, one would posit that the more specific the information that trainees can garner about their previous experience, the more likely they will be able to assess and be confident in their ability to perform the task. The relationship between improved task performance and team-efficacy has been replicated consistently in the extant literature (Brown, 2001; Brown & Sitzmann, 2011). Consonant with this perspective, Villado and Arthur (2012) found that trainees who participated in an AAR demonstrated higher levels of team-efficacy than those that did not participate in an AAR. Thus, given the large body of literature that speaks to the positive relationship between team-efficacy and performance (e.g., Arthur et al., 2007; Gully et al., 2002; Lindsley et al., 1995; Porter, 2005) as well as the results from Villado and Arthur (2012) which indicate a positive relationship between AAR and team-efficacy, it is expected that teams that demonstrate higher performance will similarly demonstrate higher levels of team-efficacy. Thus, it is hypothesized that:

H2(a): Performance will be positively related to team-efficacy.

Moreover, there will be a significant main effect for geographic dispersion and AAR conditions, such that:

H2(b): Teams in the co-located and distributed AAR conditions will have higher perceptions of team-efficacy than teams in the co-located and distributed non-AAR conditions.

H2(c): Teams in the co-located AAR conditions will have higher perceptions of team-efficacy than teams in the distributed AAR conditions.

H2(d): Teams in the co-located and distributed objective AAR conditions will have higher perceptions of team-efficacy than teams in the co-located and distributed subjective AAR conditions.

Declarative Knowledge

Declarative knowledge is the factual and conceptual information that is necessary to perform a specified task (Banks & Millward, 2007), and is a prerequisite for higher order knowledge or skill development (Ackerman, 1987; Anderson, 1982). Furthermore, declarative knowledge predicts performance for both individuals ($r = .48$, $k = 10$; Hunter & Hunter, 1984; Hunter & Schmidt, 1996) and teams ($r = .29$, $k = 24$; Devine & Phillips, 2001). For instance, teams with higher mean declarative knowledge scores outperformed teams with lower mean declarative knowledge scores on a simulated business decision-making task (Devine, 1999). Because of its consistent relationship with performance, declarative knowledge is a cognitive variable that is often incorporated into the evaluation of training interventions.

H3(a): Declarative knowledge will be positively related to performance.

However, although declarative knowledge is of clear import when evaluating training interventions, there has been little research investigating the effect of the geographic dispersion and objectivity of the AAR on the acquisition of declarative knowledge. As such, the present study sought to investigate the relationship between the different AAR conditions and the acquisition of declarative knowledge. The cognitive perspective of cooperative learning would emphasize that the effect of cognitive elaboration—active processing of information driven by team interaction—is the potential salient factor that explains the positive influence of knowledge sharing in teams (Strauss & Olivera, 2000). For example, research suggests that team experience—defined as team training—is capable of influencing the acquisition of individual knowledge (Littlepage, Robinson, & Reddington, 1997). In addition, Olivera and Strauss (2004) conducted an experiment testing the ability of teams to transfer knowledge to the individual and found that participants who worked in teams received benefits from the transfer of learning when they were re-tested using a different task as individuals. Thus, cooperative learning can influence the knowledge of the team as well as the individuals who make up that team and subsequently positively influence team performance.

It is important to note, that although it is expected that there will be differences in declarative knowledge between conditions, the content of the AAR used in the present study is not specifically designed to improve declarative knowledge. The purpose of the AAR is to allow trainees to evaluate past performance to facilitate their future performance with an emphasis on skill acquisition. However, because they were given a forum to discuss their mission performance and ways to improve in the future, it is not unreasonable to posit that the discussion would have some elements of declarative knowledge.

Although declarative knowledge has received little attention in the distributed training literature, research on virtual teams and technology-mediated interaction provides substantial evidence on the relationship between distributed communication and knowledge acquisition. Previous research indicates that many of the media used in technology-mediated interactions reduce the amount of elaboration and engender a loss in expressive and back-channel cues which subsequently reduce the effectiveness of cooperative learning (Fiore, Salas, Cuevas, & Bowers, 2003; Mesmer-Magnus, DeChurch, Jiminez-Rodriguez, Wildman, & Shuffler, 2011; Strauss & Olivera, 2000).

For instance, Baltes, Dickson, Sherman, Bauer, and LaGanke's (2002) meta-analysis indicated that not only did technology-mediated interaction have a negative influence on team effectiveness (d = -.40, k = 24), but that effect was exacerbated when the task was a knowledge-based task (d = -.50, k = 16). Thus, it would seem that the reduction in communication cues inherent in technology-mediated interaction might hinder the ability of team members or trainees to properly exchange pertinent task-related information. In addition, Webb (1992) posited that the effectiveness of knowledge acquisition through team interaction was a function of two factors, specifically, the richness of the discussion as well as the ability to apply the topics of discussion. Thus, given the reduced information richness inherent in distributed communication, it is posited that there will be a significant main effect for geographic dispersion and AAR conditions, such that:

H3(b): Teams in the co-located and distributed AAR conditions will have higher declarative knowledge scores than teams in the co-located and distributed non-AAR conditions.

H3(c): Teams in the co-located AAR conditions will have higher declarative knowledge scores than teams in the distributed AAR conditions.

Consonant with the reasoning that increased accuracy in feedback would lead to performance gains, an argument could be made for the effect of accuracy on knowledge acquisition. The performance appraisal literature has provided a large body of evidence for the effectiveness of accurate feedback (Denisi & Sonesh, 2011). Specifically, more objective feedback shoud allow trainees to identify more critical incidents and thus provide more descriptive feedback (DeNisi et al., 1989; Prince et al., 2005). That is, errors in recall, related to subjective interpretations would likely influence the effectiveness of recalled events, thus inhibiting the ability of team members to gather task-relevant information. Hence, it is posited that increases in objectivity would benefit knowledge acquisition through a richer discussion and more accurate assessments of previous performance. Thus, it is expected that there will be a significant main effect for the type of AAR review, such that:

> *H3(d):* Teams in the co-located and distributed objective AAR conditions will demonstrate higher declarative knowledge scores than teams in the co-located and distributed subjective AAR conditions.

Team Voice

Team communication refers to team members utilizing information sharing techniques to exchange information (Stevens & Campion, 1994). Communication is conceptualized as a generic teamwork competency that is required in all interdependent team-based tasks and jobs. The main function of communication is to provide a mechanism by which team members can coordinate their actions (Cannon-Bowers, Salas, Tannenbaum, & Mathieu, 1995; Ellis et al., 2005; Marks, Zaccaro, & Mathieu, 2000). The extant literature typically indicates that teams that overtly communicate more frequently outperform teams that communicate less frequently (e.g., Bowers, Jentsch, Salas, & Braun, 1998; Cannon-Bowers & Salas, 1997).

However, the focus of the present study is not on the amount of communication but rather the extent to which team members felt they were free to express their opinions. Van Dyne and Lepine (1998) define team voice as the extent to which individuals communicate freely within their team. Team voice has demonstrated positive relationships with supervisors' ratings of performance over and above non-discretionary behaviors (Van Dyne & Lepine, 1998). Thus, by providing a structured forum to facilitate communication (the AAR) in which all participants are able to discuss their ideas, trainees should feel more comfortable expressing their opinion regarding previous team performance and future action steps. Because the AAR is defined as a discussion forum in which the teams learn what happened, why it happened, and how to sustain strengths and improve on weaknesses, with the underlying premise that each individual's perspective is important, one would expect AAR team members to be more likely to express their opinions than teams that lack such a structure. Hence, one would expect that when team members feel more comfortable expressing their opinions, more strategies will be assessed and teams will be more likely to identify an effective strategy (Detert & Burris, 2007; Lepine & Van Dyne, 2001; Whiting, Podsakoff, & Peirce, 2008). For example, Whiting et al. (2008) found that voice behaviors were positively related to supervisor ratings of performance. They posit that team voice allows people to present constructive suggestions that improve perceptions of and actual performance on the task. Hence, it was posited that:

H4(a): Team voice will be positively related to performance.

Not surprisingly, there is a large body of literature demonstrating that co-located teams are more likely to effectively communicate and share information in a timely manner (Allen, 1977; Ancona & Caldwell, 1992; Bordia, 1997; Brown & Eisenhardt, 1995). However, the effectiveness of communication is commonly moderated by the extent to which team members perceive that they have an equal voice in decisions. For instance, distributed environments may represent a situation in which more team members feel comfortable expressing their thoughts and ideas because of the higher levels of anonymity associated with distributed training environments. For example, Sia, Tan, and Wei (2002) found that by increasing anonymity through the removal of visual cues, there was an increase in team polarizing communication, that is, the tendency to develop more extreme thoughts following a team discussion.

In addition, Triana, Kirkman, and Wagstaff (2011) found that reducing the salience of the minority team members' identity (e.g., females) by having teams interact in computer-mediated communication prior to face-to-face communication, led to the minority team members feeling more included in the team. That is, the type of communication medium is not the only factor that led to feelings of inclusion, the order of communication was also important. Thus, situations where teams perform in a co-located setting, but are able to train in a distributed setting may positively influence minority and majority team members' perceptions of equity and inclusion in performance situations. However, whereas minority team members felt more included, there was no difference in the extent to which minority members actually communicated. Thus, feeling more included or having higher perceptions of team voice did not manifest itself in more actual communication.

The social model of deindividuation effects postulates that an individual's proclivity to express their opinion is a function of their perceptions of anonymity (Spears & Lea, 1994). Thus, because technology-mediated interaction leads to greater perceived anonymity, individuals are able to "hide behind" the technology and express their opinions seemingly free of consequence. For example, Jessup, Connolly, and Galegher (1990) found that teams with greater anonymity not only provided more solutions to an idea-generation problem, but they also demonstrated more critical comments than identified teams. Consequently, it would seem that the increased anonymity that results from distributed training environments, should translate into trainees feeling more comfortable about expressing their thoughts and ideas freely. Hence, it is posited that there will be a significant main effect for geographic dispersion (in the AAR conditions), such that:

H4(b): Teams in the distributed AAR conditions will have higher perceptions of team voice than teams in the co-located AAR conditions.

However, there is no reason to expect that the objectivity of the AAR will differentially impact the trainees' perceptions of their ability to communicate freely.

H4(c): Team members' perceptions of team voice for teams in the co-located and distributed subjective AAR conditions and teams in the co-located and distributed objective AAR conditions are not expected to be different.

Team Cohesion

Team cohesion is classically defined as "the resultant forces that are acting on the members to stay in a group" (Festinger, 1950, p. 274). Barrick, Stewart, Neubert, and Mount (1998) describe team cohesion as "synergistic interactions between team members, including positive communication, conflict resolution, and effective workload sharing" (p. 382). Team cohesion is considered to be an important teamwork process variable because of its positive relationship with team performance. Meta-analytic evidence shows a moderate relationship between team cohesion and team performance (d = 0.30, k = 19) in terms of behavioral outcomes (Beal, Cohen, Burke, & Mclendon, 2003). However, the relationship is weaker when team performance is operationalized in terms of objective outcomes (d = 0.17, k = 47; Beal et al., 2003). Task interdependence also influences the relationship between team cohesion and performance, such that performance on tasks with greater interdependence demonstrates a stronger relationship with team cohesion (Gully, Devine, & Whitney, 1995).

H5(a): Team cohesion will be positively related to performance.

Given the definition posed by Barrick et al. (1998), it would seem that the AAR would provide a structured and consistent opportunity for teams to enact these "synergistic interactions." That is, the AAR allows individuals the opportunity to determine what positive actions and interactions they can draw from previous performance (positive communication), potential negative actions and their potential causes (conflict resolution), and finally the opportunity to strategize about their future team performance from their lessons learned (potentially workload sharing). In addition, a negative relationship between team cohesion and role ambiguity has been reported, such that individuals who are uncertain about the scope of their responsibility experienced lower perceptions of team cohesion (Eys & Carron, 2001). As such, AAR teams who are able to discuss their performance and their action steps for future performance sessions should have lower role ambiguity and subsequently higher levels of team cohesion.

The relationship between team cohesion and performance (Beal et al., 2003; Gully et al., 1995) and the importance of team viability, provides the basis for the interest in team cohesion when evaluating team training interventions. Within the context of the present study, one would posit that team cohesion would be affected by the geographic dispersion of trainees. Distributed team members may have difficulty developing strong normative bonds given their reduced ability to feel part of the team and the increased levels of deindividuation via technology-mediated communication (Spears & Lea, 1994). Specifically, distributed training protocols are likely to lead to greater anonymity between participants (Jessup et al., 1990) and reduced trust between trainees (Kring, 2004; Rosen, Furst, & Blackburn, 2006); factors that are likely to affect the formation of team cohesion. For example, in comparisons of distributed and co-located teams, researchers commonly find that distributed teams demonstrated lower levels of team cohesion in both lab (Bouas & Arrow, 1996; Warkentin, Sayeed, & Hightower, 1997) and field studies (Fjermestad & Hiltz, 1999). Thus, there will be a significant main effect for geographic dispersion and AAR conditions, such that:

H5(*b*)*:* Teams in the co-located and distributed AAR conditions will have higher perceptions of team cohesion than teams in the co-located and distributed non-AAR conditions.

H3(c): Teams in the co-located AAR conditions will have higher perceptions of team cohesion than teams in the distributed AAR conditions.

H5(d): Team members' perceptions of team cohesion for teams in the co-located and distributed subjective AAR conditions and teams in the co-located and distributed objective AAR conditions are not expected to be different.

Team-Level Reactions

The effectiveness of any training intervention is determined by its influence on the training evaluation outcomes of interest. Kirkpatrick (1959, 1976, 1996) proposed a four-level model of training evaluation that continues to be the most popular to date. Specifically, the four training evaluation criteria outlined in Kirkpatrick's (1959, 1976, 1996) model are reaction, learning, behavior, and results. Reaction criteria are the most commonly used training evaluation criteria and are intended to measure the extent to which trainees liked the training (Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997). Learning criteria are measures of the amount of knowledge acquisition that takes place; whereas behavioral criteria are the effects the training has on job-related behaviors (Arthur et al., 2003). Finally, results criteria are more macro indicators of training effectiveness and typically focus on the impact of training at the organizational level. This sometimes takes the form of applying utility analysis to quantify the effectiveness of training in monetary terms. It is important to note that a common misconception of this model is that it is hierarchical, such that the training must demonstrate one level to achieve a "higher" level. While this may be true for some levels (e.g., learning and behavior), it is not necessary for a training to demonstrate positive reactions in order for it to demonstrate positive outcomes at the other levels.

Kraiger, Ford, and Salas (1993) extended Kirkpatrick's model positing that there is both a learning and affective component to reaction criteria. Learning or utility reactions are the extent to which trainees' perceive the training to be useful and affective reactions are how much the trainees enjoyed the training. Affective and utility reactions are commonly measured at the individual level and post-training. To this author's knowledge there have not been any studies that have investigated affective and utility reactions at the team level. However, given that the present study's AAR is a team-level intervention and all of the outcomes are at the team level, the focus of reactions is at the team level, which is a unique feature of this study.

Arthur et al. (2003) found that organizational training interventions are commonly rated positively by trainees (d = .60, k = 15). Although, individual reaction criteria have demonstrated a minimal correlation with actual individual learning (ρ = .02, k = 11) and behavior (ρ = .03, k = 9) outcomes (Alliger et al., 1997), they continue to be a critical outcome variable in training evaluation. One potential reason for the continued use of reactions in training evaluation is that they are positively related to trainee motivation and self-efficacy (Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008). In addition, formal training often serves a socialization function, beyond its principle purpose of disseminating knowledge. Thus, to the extent that trainees have favorable perceptions or reactions to a training intervention, the more likely it is to achieve its socialization goals (Feldman, 1989; Kraiger et al., 1993). Finally, in organizations where training is frequently required and the information in that training is critical to performance, positive affective reactions may increase the likelihood that employees will attend the necessary training modules. That is, when trainees have negative reactions to the training, it may decrease the likelihood that they will participate in future training sessions, pay attention while in those sessions, or transfer the training content to their jobs. Thus, although reaction criteria are not necessarily related to learning or behavioral outcomes, they remain an important training evaluation outcome of interest.

A detailed search failed to identify any work investigating trainee (or team-level) reactions to distributed training, and specifically, AAR training. As such, it was deemed important to investigate the extent to which team-level reactions to the training were influenced by the objectivity of the AAR and geographic dispersion. Thus, to this author’s knowledge this study will provide the first examination of team-level reactions and their relationship with other team-level training outcomes. Consequently, the present study sought to answer the following research question:

> *Research Question:* How will the different AAR training conditions influence the teams’ reactions to the AAR training?

A summary of the present study’s hypotheses is presented in Table 1.

Table 1
Research Hypotheses

Hypothesis		Support
H1(a)	Teams in the co-located and distributed AAR condition will have higher performance scores than teams in the co-located and distributed non-AAR condition.	Yes
H1(b)	Teams in the co-located AAR conditions will have higher performance scores than teams in the distributed AAR conditions.	No
H1(c)	Teams in the co-located and distributed objective AAR conditions will display higher performance than teams in the co-located and distributed subjective AAR conditions.	No
H2(a)	Performance will be positively related to team-efficacy.	Yes
H2(b)	Teams in the co-located and distributed AAR conditions will have higher perceptions of team-efficacy than teams in the co-located and distributed non-AAR conditions.	Yes
H2(c)	Teams in the co-located AAR conditions will have higher perceptions of team-efficacy than teams in the distributed AAR conditions.	No
H2(d)	Teams in the co-located and distributed objective AAR conditions will have higher perceptions of team-efficacy than teams in the co-located and distributed subjective AAR conditions.	No
H3(a)	Declarative knowledge will be positively related to performance.	Yes
H3(b)	Teams in the co-located and distributed AAR conditions will have higher declarative knowledge scores than teams in the co-located and distributed non-AAR conditions.	No
H3(c)	Teams in the co-located AAR conditions will have higher declarative knowledge scores than teams in the distributed AAR conditions.	No
H3(d)	Teams in the co-located and distributed objective AAR conditions will demonstrate higher declarative knowledge scores than teams in the co-located and distributed subjective AAR conditions.	No
H4(a)	Team voice will be positively related to performance.	Yes
H4(b)	Teams in the distributed AAR conditions will have higher perceptions of team voice than teams in the co-located AAR conditions.	No
H4(c)	Team members' perceptions of team voice for teams in the co-located and distributed subjective AAR conditions and teams in the co-located and distributed objective AAR conditions are not expected to be different.	No
H5(a)	Team cohesion will be positively related to performance.	Yes
H5(b)	Teams in the co-located and distributed AAR conditions will have higher perceptions of team cohesion than teams in the co-located and distributed non-AAR conditions.	Yes
H3(c)	Teams in the co-located AAR conditions will have higher perceptions of team cohesion than teams in the distributed AAR conditions.	No
H5(d)	Team members' perceptions of team cohesion for teams in the co-located and distributed subjective AAR conditions and teams in the co-located and distributed objective AAR conditions are not expected to be different.	No
Research Question: How will the different AAR training conditions influence the teams' reactions to the AAR training?		—

Note. "Yes" indicates there was support for the hypothesis. "No" indicates the hypothesis was not supported.

CHAPTER II

METHOD

Participants

Participants were recruited from the human subject pool of Texas A&M University's psychology department. The sample consisted of 492 individuals (47% female) who participated in 123 4-person teams. Participants reported a mean age of 18.84 years (*SD* = 1.21) and described themselves as having average video-game experience (*M* = 1.81, *SD* = 0.65; video-game experience was measured using a 3-point scale where 1 = novice, 2 = average, and 3 = expert). Participants were provided with course credit for their participation. Additionally, to motivate them to remain focused and attempt to improve their performance during the study, participants in the first, second, and third highest performing teams in each of the six conditions were awarded $80, $40, and $20, respectively. Overall and condition-specific demographic information for the study participants are presented in Table 2.

Design

This study utilized a 3 (type of AAR review: non-AAR versus subjective AAR versus objective AAR) × 2 (geographic dispersion: co-located versus distributed) × 3 (sessions) repeated measures design. Type of AAR review and geographic dispersion served as the between-subjects independent variables, and session served as the repeated or within-subjects independent variable. Six separate dependent variables—performance, declarative knowledge, team-efficacy, team voice, team cohesion, and team-level reactions—were measured periodically throughout the study protocol.

Table 2

Demographic Composition of the Sample by Training Condition

	Training Condition													
	Co-located						**Distributed**							
	Non-AAR (*n* = 23 teams)		**Subjective AAR (*n* = 20 teams)**		**Objective AAR (*n* = 20 teams)**		**Non-AAR (*n* = 20 teams)**		**Subjective AAR (*n* = 20 teams)**		**Objective AAR (*n* = 20 teams)**		**Overall**	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Sex														
Female	49	53.26	42	52.50	40	50.63	40	50.00	29	36.25	34	42.50	234	47.66
Male	43	46.74	38	47.50	39[a]	49.37	40	50.00	51	63.75	46	57.50	257	52.34
Number of Males per Team														
0	2	8.70	2	10.00	2	10.00	1	5.00	0	0.00	0	0.00	7	5.69
1	7	30.43	6	30.00	4	20.00	6	30.00	0	0.00	5	25.00	28	22.76
2	8	34.78	6	30.00	7	35.00	7	35.00	10	50.00	7	35.00	45	36.59
3	4	17.39	4	20.00	6	30.00	4	20.00	9	45.00	5	20.00	32	26.02
4	2	8.70	2	10.00	1	5.00	2	10.00	1	5.00	3	15.00	11	8.94
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Age	18.60	0.44	18.73	0.51	18.56	0.40	18.86	0.51	19.13	1.04	19.18	0.76	18.84	1.21
Video-game experience	1.91	0.71	1.76	0.56	1.81	0.70	1.76	0.56	1.83	0.66	1.75	0.61	1.81	0.65

Note. AAR = after-action review; [a]Demographic data for one participant in the co-located objective AAR training condition was missing.

Measures

Performance Task—Steel Beasts Pro PE ver. 2.370

Steel Beasts Pro PE (eSim Games, 2007) is a cognitively complex, PC-based tank synthetic task environment, allowing multiple players to be networked together to cooperatively complete missions in a simulated battlefield environment. The simulator uses highly accurate replicas of U.S. M1A1 and Russian T-72 tanks to simulate an armored warfare environment. Figure 2 presents a screenshot of the replica M1A1 tank in the simulated environment.

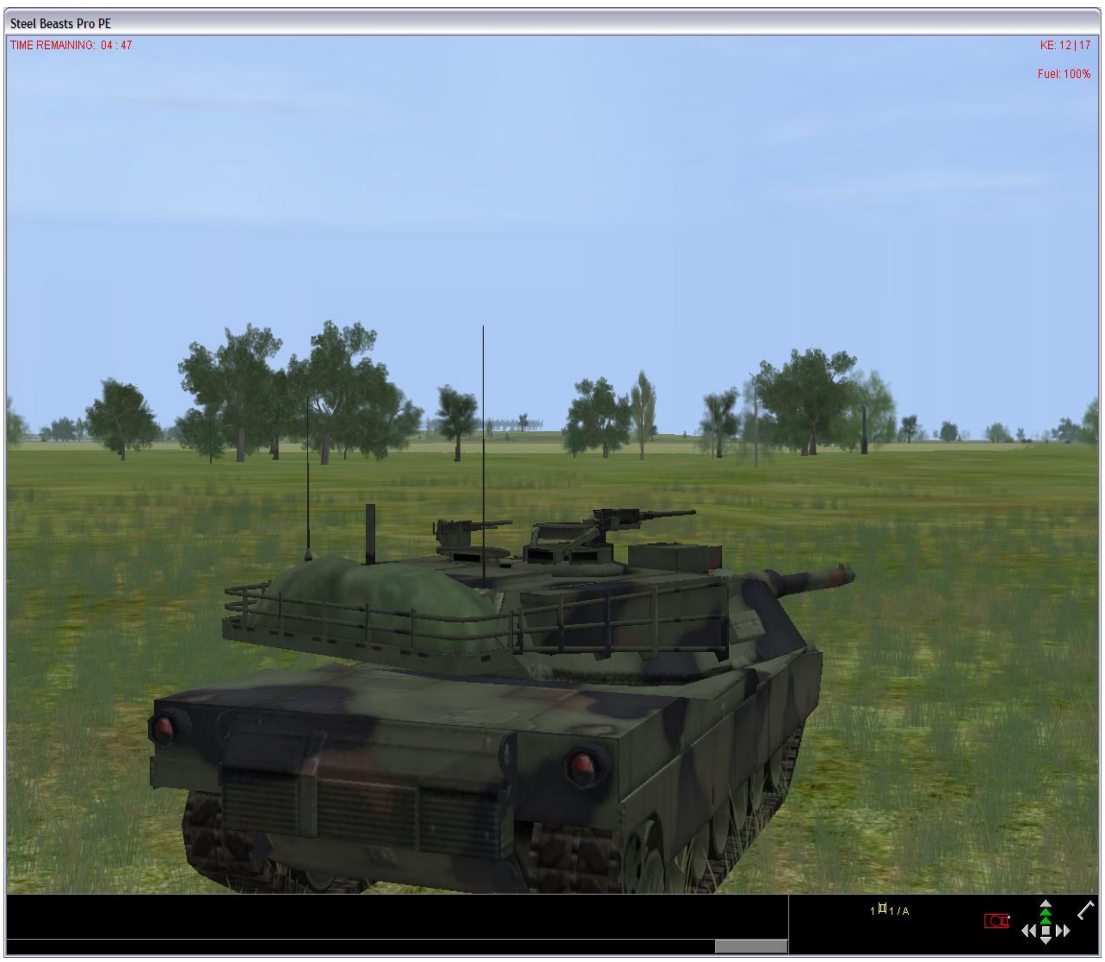

Figure 2. Steel Beasts replica M1A1 tank.

Participants operated the PC-based simulator using a monitor, keyboard, mouse, and joystick. The simulated environment consisted of a two-tank platoon of U.S. M1A1 tanks controlled by the participants. Four networked computers were used to operate the two-tank platoon; each participant had his/her own computer. Each tank in the platoon was operated by 2 participants; one participant served as the gunner and a second participant served as the commander/driver of the tank. Therefore, each team was comprised of two gunners and two commander/drivers (see Figure 3). Team members communicated with each other via voice-activated microphones and headphones.

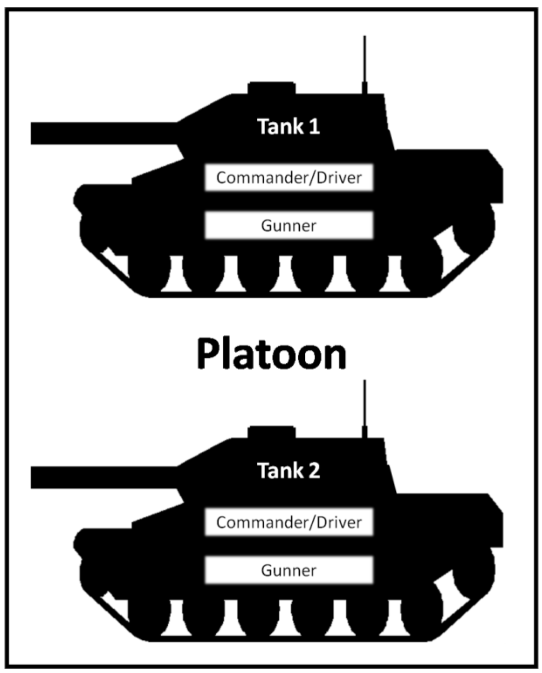

Figure 3. Illustration of the roles within and between tanks.

Multiple first-person perspective views were available to each participant, depending on their role. For example, gunners were able to switch between multiple gun sight views and a map view of the battlefield. Commander/drivers were able to switch between several views ranging from sitting inside the tank to standing up through the hatch of the tank, in addition to a view of the gunner's gun sight, and a map view of the battlefield.

The performance task was highly interdependent, with elements of both task and outcome interdependency. Task interdependency existed at the level of the tank such that the tank could not be operated successfully without the combined effort of the gunner and commander/driver. Outcome interdependency existed at the level of the team. Specifically, missions were designed such that a single tank was not able to complete the mission without the assistance of the other tank. The results of a team task analysis (Arthur, Edwards, Bell, Villado, & Bennett, 2005) confirmed that the roles were interdependent within and between tanks. The mean for the team-relatedness ratings (on a 1-5 point scale) was 4.10 (SD = 0.44) and the team-workflow rating was 4.53 (SD = 0.43) demonstrating that the task was highly team-related (i.e., *amount* of teamness) with the optimal workflow (i.e., *type* of teamness) classified as intensive interdependence. Intensive interdependence is represented by work entering into the team and team members having to diagnose, problem solve, and collaborate as a team. There is no uniform direction in which the work flows instead all team members coordinate with each to complete the task (Arthur et al., 2005, 2012).

Steel Beasts Pro PE Missions

There were two test missions in each of the three sessions. The same mission map was used for all six test missions (see Figure 4). Participants also completed two practice missions, one during Session 2 and the second during Session 3. The practice missions were identical to the test missions; however, the participants were given 15 minutes to perform the mission and were told that their scores on the practice missions would not count towards their performance scores. Each mission required a team to destroy 10 enemy tanks while the participants were en route to a target destination. Missions (both practice and test) were preceded by a 2-minute briefing and planning session during which teams were shown a mission briefing with information regarding mission objectives and rules, potential enemy positions, and enemy capabilities. Teams were then encouraged to formulate a strategy to complete the mission. After the briefing and planning session, for test missions, teams were allowed 10 minutes to complete the mission. A mission ended when (a) the team completed all mission objectives, (b) all participant tanks were destroyed, or (c) the 10-minute time limit expired. In contrast to the test missions, for practice missions, teams were allowed 15 minutes which they could use for either planning or interacting with the simulator. The first practice mission provided participants with suggested waypoints for optimal performance of the missions (see Figure 5), whereas the second practice mission was identical to the test missions, with the exception of the time limit.

Performance scores were obtained at the team level. Teams earned points for the number of enemy tanks destroyed (5 points per tank), and advancing beyond certain boundaries (2.5 points per tank per boundary crossed [e.g., Sierra and Alpha in Figure 4] and 12.5 for each tank that reached the objective). And teams lost points for destroying one of their own tanks (fratricide; -50 points). Thus, the total possible points ranged from -50 to +100. As previously noted, team performance for each session was operationalized as the average of the team's scores for the two test missions that were performed in each session.

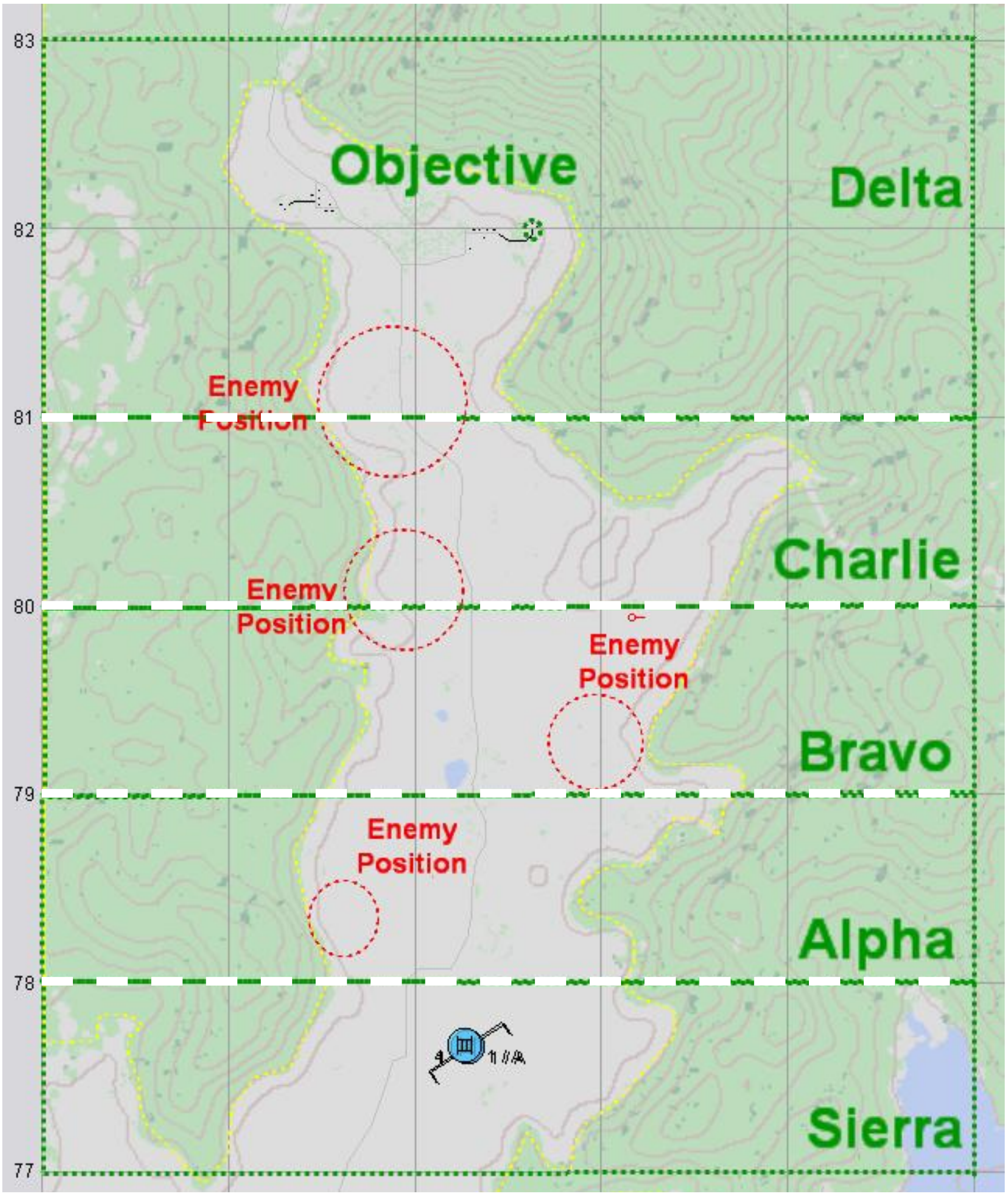

Figure 4. Mission map for test missions and the second practice mission.

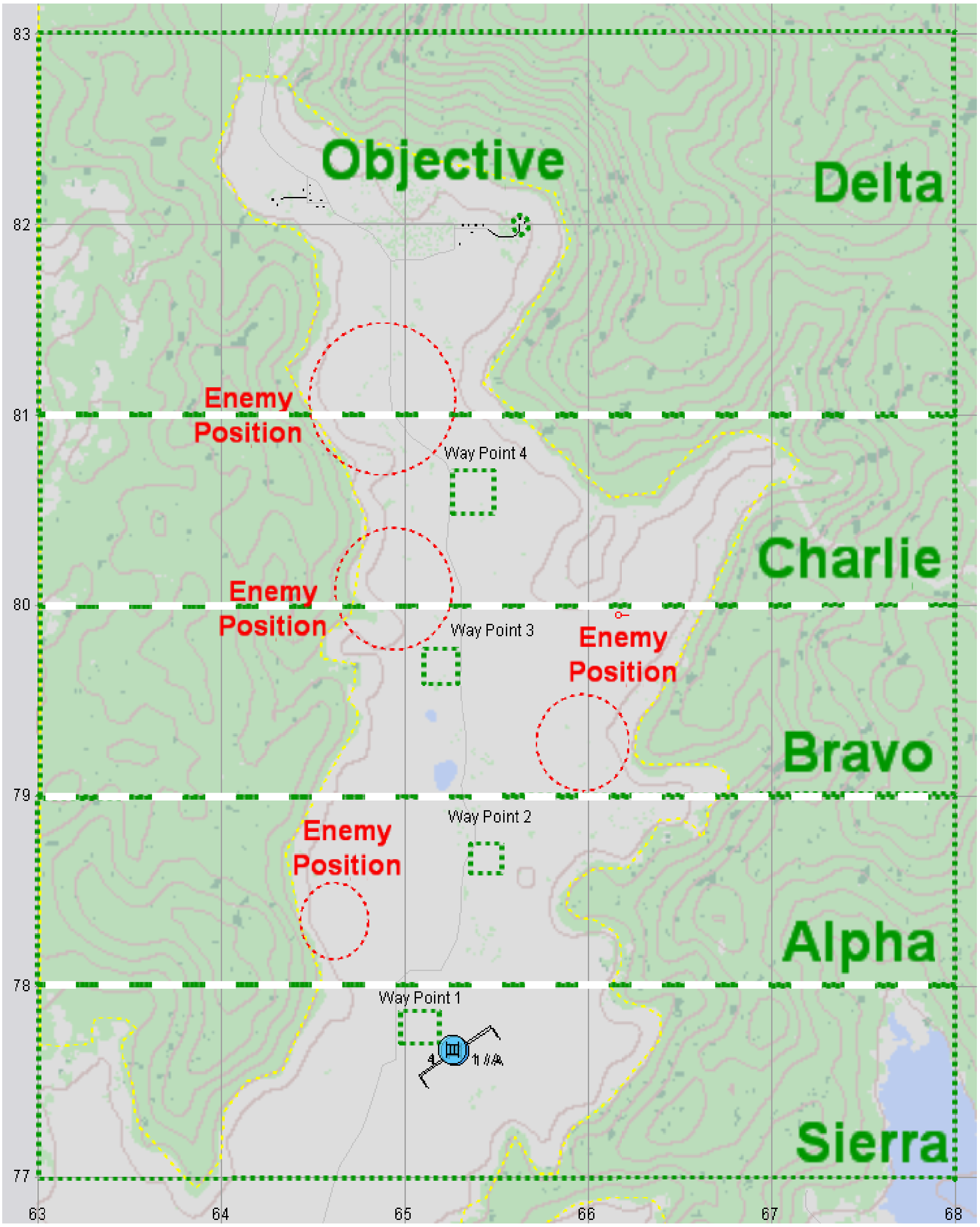

Figure 5. Mission map for first practice mission.

Overall (total) performance was operationalized as the average performance across all three sessions. The method used to determine performance scores was explained to participants during each mission briefing and performance scores were available for them to review at the conclusion of each mission.

Declarative Knowledge

Declarative knowledge was assessed using a 30-item, 3-alternative multiple-choice test. This test was developed using Arthur, Edwards, Bell, and Bennett's (2002) test as a model. A test-retest reliability of .89 (N = 240; time interval = 4 hours) has been recorded for scores on this test (Villado, 2008). The declarative knowledge measure is presented in Appendix A.

Team-Efficacy

A modified version of the Arthur et al. (2007) measure was used to assess team-efficacy. The measure consisted of six task-specific items with a team referent. Participants provided their ratings using a 5-point Likert rating scale (1 = strongly disagree, 5 = strongly agree). Team-efficacy scores were calculated using the average of the mean individual-level item responses. Internal consistency estimates for the first and second administrations of the team-efficacy scores at the individual-level of analysis were .92 and.93, respectively (N = 492). These estimates are slightly higher than those reported by Arthur et al. (2007) who obtained a range of internal consistencies between .76-.84 across multiple administrations. The team-efficacy measure is presented in Appendix B.

Team Voice and Team Cohesion

Team voice and team cohesion were assessed using a 10-item measure that consisted of 4 team voice and 6 team cohesion items. The 10 items were selected from Barry and Stewart's (1997) Group Process measure, and modified to fit the present performance task. Participants provided their ratings using a 5-point Likert rating scale (1 = strongly disagree, 5 = strongly agree). The internal consistency estimate for the team voice scores at the individual-level of analysis (N = 492) was .72; and it was .87 for the team cohesion scores. Team voice and team cohesion scores were calculated using the average of the mean individual-level item responses for each subscale. The team voice and team cohesion measures are presented in Appendix C.

Demographics

Participants reported their age, sex, experience with video games, and whether they had previous experience with Steel Beasts. A single video-game experience item asked participants to describe their general experience with video games using a 3-point scale (i.e., where 1 = novice, 2 = average, 3 = expert). Prior experience with Steel Beasts was collected with the intention of eliminating participants who had prior experience with the task. However, no participant reported any prior experience with Steel Beasts and so no one was removed from the study for this reason. The demographics measure is presented in Appendix D. For a breakdown of sex per condition and team along with other demographic information see Table 2.

Reactions to the Training

Reactions to the AAR were measured using 12 items, six of which assessed affective reactions (e.g., Participating in the after-action-review mades me feel uncomfortable), and six assessed utility reactions (e.g., Participating in the after-action-review helped me learn to play *Steel Beasts*). Thus, reactions to the AAR referred to the affective reactions associated with participating in the AARs and the perceptions of the effectiveness of the AAR. Ratings were made on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Reactions to the AAR were measured after the final mission (i.e., second test mission of Session 3). The measure was scored by computing the mean of the 12 items across all four team members. The internal consistency for reactions to the AAR ratings was .89 (N = 316) at the individual level. The trainee reactions measure is presented in Appendix E.

Training Manipulation

Participants were trained to operate the simulator first as individuals and then as a team. During the individual training phase, participants were allowed 45 minutes to complete nine training tutorials. Each tutorial began with participants reading the tutorial content from a tutorial handbook. Once participants understood the content and objectives of the tutorial, they then completed a mission that provided hands-on practice of the tutorial content. Subsequent tutorials continued using the same procedure. Six of the training tutorials focused on tasks relevant to a participant's role, and the remaining three tutorials focused on tasks relevant to their teammate's role.

During the performance phase, participants operated the simulator as a team to complete the six test missions. All participants completed the same test missions, regardless of training condition. The events that followed each team performance mission depended on the training condition to which the team was assigned.

Within the co-located training condition, teams were randomly assigned to a non-AAR, subjective AAR, or objective AAR condition. Teams assigned to the distributed training condition were also randomly assigned to a non-AAR, subjective AAR, or objective AAR training condition. However, in contrast to the co-located training condition, participants on each team assigned to the distributed training condition were separated geographically. Specifically, the two participants assigned to Tank 1 remained in the co-located laboratory (henceforth referred to as the main laboratory; Figure 6) and the individuals in Tank 2 were escorted to a laboratory (henceforth referred to as the distributed laboratory) in a separate building on campus to participate in the study (see Figure 7). The distributed laboratory had the same equipment as the co-located laboratory. That is, participants in the distributed laboratory operated Steel Beasts using the same monitors, keyboards, mice, and joysticks as their teammates in the main laboratory and communicated with each other and their teammates in the main laboratory via voice activated microphones and headphones.

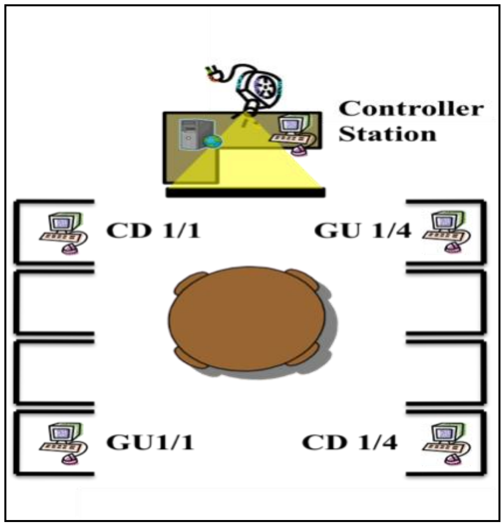

Figure 6. Co-located lab configuration. CD = commander/driver; GU = gunner. 1/1 and 1/4 represent the tank designations.

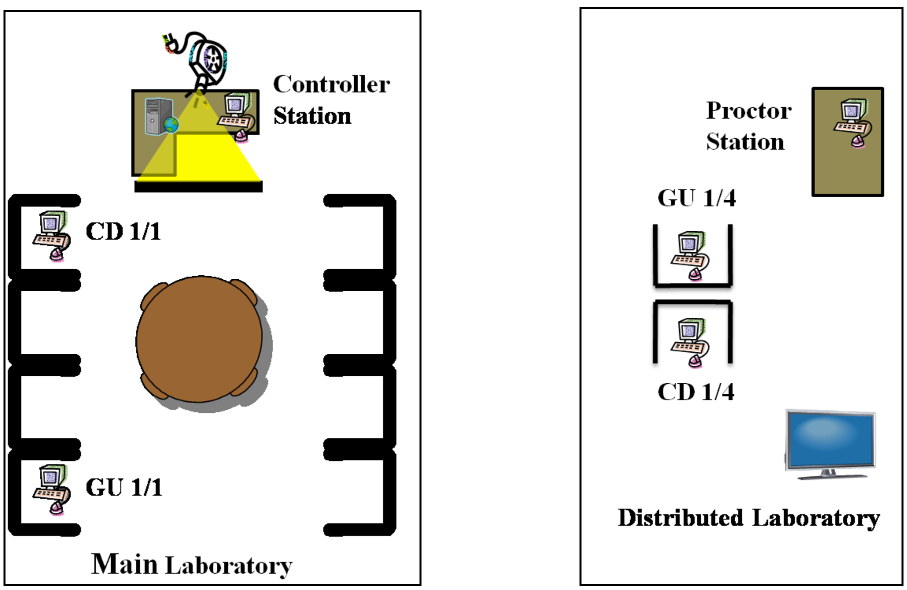

Figure 7. Distributed lab configuration. CD = commander/gunner, GU = gunner. 1/1 and 1/4 represent the tank designations.

Non-AAR Training Conditions

Once a test mission ended, participants assigned to the co-located non-AAR training condition completed a filler task that was unrelated to the Steel Beasts. This was to give a similar break between missions for the non-AAR and AAR training conditions.

Subjective AAR Training Conditions

After the two practice sessions and after test missions 2 and 4, participants in the subjective AAR training condition participated in a 10-minute AAR, monitored by a facilitator. Prior to the first AAR, the facilitator explained the AAR process to team members and provided teams with a form that detailed each step of the AAR process. After introducing participants to the AAR process, facilitators only intervened during AARs to ensure that teams completed each step of the AAR in the order presented in Figure 1 within the specified time limits.

Subjective AARs began with participants recalling the intended outcome and the actual outcome of their most recently completed mission. Participants then compared the two to determine whether their goals had been met. Next, participants identified specific behaviors or events that contributed to or detracted from achieving the mission objectives. The participants were then encouraged to set specific and difficult, yet attainable goals for the subsequent mission. Each AAR concluded with participants identifying behaviors and actions that would increase the likelihood of meeting their self-set goals and subsequent mission objectives. Teams then completed the specified paper-and-pencil measures as per Table 3.

For the distributed condition each AAR was guided by a facilitator who was located in the main laboratory. Each AAR was conducted in the same manner as the co-located subjective AARs, except that participants communicated via voice-activated microphones and headphones rather than in the face-to-face manner described in co-located subjective AAR. Similar to the co-located condition, once they had completed the AAR, participants then completed any measures scheduled to follow that mission or if there were no measures they immediately logged back into the simulator.

Objective AAR Training Conditions

Participants assigned to the objective AAR training condition participated in an AAR after the two practice sessions and after test missions 2 and 4 (and within the same 10-minute time period as participants in the subjective AAR training condition). However, participants in the objective AAR training condition reviewed the progress of their most recently completed mission using the simulator's review tool, operated by the facilitator. The review tool allowed participants to replay, pause, and move forward or backward through the simulated environment of the most recently completed mission and was displayed on a projector to all of the participants. Participants could view the mission progress from multiple perspectives and examine it from any point in the simulated environment (e.g., from either tanks' perspective, the enemy's perspective, or a top-down view of the mission). They followed the same procedure as the teams in the co-located condition. After the AAR, teams completed the specified paper-and-pencil measures as per Table 3.

For the distributed objective AAR condition, each AAR was conducted in the same manner as the co-located objective AARs, except that participants communicated via voice-activated microphones and headphones rather than face-to-face. In addition, each laboratory had a monitor that provided the same video playback simultaneously to the all four participants (two in the main lab and two in the distributed lab). Similar to the co-located condition, once they had completed the AAR, participants then completed any measures scheduled to follow that mission or if there were no measures they immediately logged back into the simulator.

Procedure

The study protocol lasted five hours and was divided into three phases. Participants were randomly assigned to a specific role within the team, either the gunner or commander/driver position of the performance task, as well as a specific tank. Prior to the first phase individuals in the distributed condition were separated by tank with two individuals in the main laboratory and two individuals in the distributed laboratory. An effort was taken to minimize the amount of communication between the trainees prior to the start of the experiment. Two proctors led the distributed participants through the first phase independently.

Table 3

Schedule of Activities for Each Training Session by Training Condition

Session	**Scheduled Activities**					
0	Informed consent Declarative knowledge [1] Video-game experience Demographics Individual tutorials					
	Training Conditions					
	Co-located			**Distributed**		
	Non-AAR (*n* = 23 teams)	**Subjective AAR (*n* = 20 teams)**	**Objective AAR (*n* = 20 teams)**	**Non-AAR (*n* = 20 teams)**	**Subjective AAR (*n* = 20 teams)**	**Objective AAR (*n* = 20 teams)**
1	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] Filler Task Team-efficacy [1]	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] AAR Team-efficacy [1]	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] AAR Team-efficacy [1]	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] Filler Task Team-efficacy [1]	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] AAR Team-efficacy [1]	Planning Test Mission 1 [Bl1] Planning Test Mission 2 [Bl2] AAR Team-efficacy [1]
2	Planning Practice Mission 1 Filler Task	Planning Practice Mission 1 AAR	Planning Practice Mission 1 AAR	Planning Practice Mission 1 Filler Task	Planning Practice Mission 1 AAR	Planning Practice Mission 1 AAR
	Planning Test Mission 3 Planning Test Mission 4 Filler Task	Planning Test Mission 3 Planning Test Mission 4 AAR	Planning Test Mission 3 Planning Test Mission 4 AAR	Planning Test Mission 3 Planning Test Mission 4 Filler Task	Planning Test Mission 3 Planning Test Mission 4 AAR	Planning Test Mission 3 Planning Test Mission 4 AAR
3	Planning Practice Mission 2 Filler Task	Planning Practice Mission 2 AAR	Planning Practice Mission 2 AAR	Planning Practice Mission 2 Filler Task	Planning Practice Mission 2 AAR	Planning Practice Mission 2 AAR
	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]	Planning Test Mission 5 Planning Test Mission 6 Team-efficacy [2]

Table 3 (cntd.)

Session	Scheduled Activities					
	Training Conditions					
	Co-located			**Distributed**		
	Non-AAR (*n* = 23 teams)	**Subjective AAR (*n* = 20 teams)**	**Objective AAR (*n* = 20 teams)**	**Non-AAR (*n* = 20 teams)**	**Subjective AAR (*n* = 20 teams)**	**Objective AAR (*n* = 20 teams)**
3 (cntd.)	Declarative knowledge [2] Team Voice Team cohesion Reaction	Declarative knowledge [2] Team voice Team cohesion Reaction	Declarative knowledge [2] Team voice Team cohesion Reaction	Declarative knowledge [2] Team voice Team cohesion Reaction	Declarative knowledge [2] Team voice Team cohesion Reaction	Declarative knowledge [2] Team voice Team cohesion Reaction

Note. AAR = after-action review. Planning periods were limited to 2 minutes, test missions were limited to 10 minutes, practice missions were limited to 15 minutes, and AARs were limited to 10 minutes. Bl1= Baseline 1; Bl2 = Baseline 2; a team's baseline score is the mean of the two baseline mission.

For the co-located condition all four participants were situated in the main laboratory. During the first phase of the study, participants were familiarized with the protocol, completed the informed consent form, and the baseline Steel Beasts declarative knowledge measure.

During the second phase of the study, participants began their individual simulation training. Participants completed the tutorials on four individual computers and monitors using the keyboard and a right-handed joystick to navigate through the tutorials/missions. The joystick controlled the participants' viewpoint and was used to judge distances and fire at enemy targets. Each computer had a headset that allowed participants to listen individually to the tutorials and later in team missions communicate with other team members. Trainees were given 45 minutes to read and complete all of the tutorials. For the first tutorial, the researcher read the tutorial to the participants as they followed along in their tutorial handbooks. After completing the first tutorial, participants then completed the remaining tutorials at their own pace. Each tutorial began with participants reading the tutorial content from a tutorial handbook. Once participants understood the content and objectives of the tutorial, they then completed a tutorial-based mission that provided hands-on practice of the tutorial content. Subsequent tutorials continued following the same procedure. Participants who completed their tutorials before the 45-minute time limit were allowed to repeat any of the tutorials if they chose to. In the distributed conditions, the two proctors communicated throughout the first and second phase in order to ensure the participants were on a similar pace. After the second phase of the study all participants received a short break.

Upon completing the tutorials, participants began the third and final phase of the protocol, the team-training phase. This phase of the protocol commenced with participants being shown how to use the voice activated microphones and headphones. At this point, all participants put on their headsets which were used as the main form of communication throughout the remainder of the session, with the exception of co-located AARs. Participants were asked to demonstrate their ability to use the microphones and headsets. Once all participants (co-located and distributed) had their headsets on and they were working properly, they were given verbal instructions by the proctor in the main laboratory to begin the first team mission.

Each team mission began with a planning period. Participants were allowed 2 minutes to review the mission briefing and map, formulate a strategy, and discuss the strategy with their teammates during the planning period. Teams were allowed to begin the mission prior to the 2-minute time limit if all team members were ready to do so and agreed to it. Otherwise, the team mission began after two minutes had expired. Teams were allowed 10 minutes to complete each team mission. The simulator displayed the mission runtime. Trainees completed the test missions in two mission blocks (see Table 3). The first two test sessions were followed by a practice session that lasted 15 minutes. The practice session scores were not counted towards the teams' performance score. In addition, if a practice mission ended early, participants were able to restart the mission and use the entire 15 minutes. Once a team completed a particular mission (see Table 3) or the mission was terminated, teams in the AAR conditions participated in the AAR process. Team mission briefing (2 minutes), team mission (10 minutes), and AAR (10 minutes) time limits were established and deemed to be sufficient on the basis of pilot testing.

CHAPTER III

RESULTS

Baseline Analyses

Prior to the analyses, the data were screened and the assumptions of the specific statistical analyses were evaluated, as recommended by Tabachnick and Fidell (2007). There were no systematic missing data and the data were adequately normally distributed. To assess the probability of detecting the effects indicated by the hypotheses, a power analysis was conducted using G*Power 3.1 (Faul, Erdfelder, Buchnar, & Lang, 2009), using the meta-analytic effect size from Schurig et al. (2011; d = 1.12), the power analysis indicated that—with the current sample of 123 teams and an alpha of .05—the conditions of the present study result in a 94% chance of detecting an effect between the AAR and non-AAR training conditions.

Once the screening procedures were completed, the individual-level data (i.e., declarative knowledge, team-efficacy, team voice, team cohesion, and reactions) were evaluated to justify aggregation to the level of the team. Agreement and reliability indices (i.e., $r_{wg(1)}$, $r_{wg(j)}$, r^*_{wg}, ICC_1, and ICC_2) were calculated to assess the appropriateness of aggregating the individual-level data to the team-level (James, Demaree, &Wolf, 1984; Lindell, Brandt, & Whitney 1999). The reliability estimates and median agreement indices suggested that aggregation to the team-level was appropriate (see Table 4). Therefore, team-level scores for declarative knowledge, team-efficacy, team voice, team cohesion, and reactions were created by averaging the individual-level scores within the 4-person teams.

Table 4
Individual-Level Study Variables Intraclass Correlation Coefficients and Median Agreement Indices

Measure	*N*	ICC_1	ICC_2	$r_{wg(1)}$	$r_{wg(j)}$	r^*_{wg}
Declarative Knowledge (30 items)	492	.09	.47	.91	—	—
Team-efficacy (6 items)	492	.15	.55	.92	.92	.68
Team Voice (4 items)	492	.17	.50	.91	.90	.66
Team Cohesion (6 items)	492	.12	.56	.88	.93	.72
Team-Level Reactions (12 items)	316	.14	.49	.92	.89	.66

Note. Median $r_{wg(1)}$ and $r_{wg(j)}$ were calculated using the formulas presented by James et al. (1984). Median r^*_{wg} was calculated using the formulas presented by Lindell et al. (1999).

Performance, declarative knowledge, and video-game experience were assessed prior to the participants interaction with the AAR training manipulation. Thus, a group of 1-way analysis of variance (ANOVA) tests, with the six different training conditions acting as the between-subjects variable, were conducted to test the effectiveness of the random-assignment into conditions across the variables of interest. The results indicated that there were no mean differences by condition for performance $F(5, 117) = 1.90$, $p > .05$, $\eta^2 = .08$, declarative knowledge $F(5, 117) = 0.76$, $p > .05$, $\eta^2 = .02$, or video-game experience $F(5, 117) = 1.65$, $p > .05$, $\eta^2 = .04$. Thus, the use of random-assignment to conditions as a means of controlling for pre-existing differences was considered successful. Consequently, video-game experience was not used as a covariate and will not be discussed further. However, given that team sex composition displayed a significant relationship with performance, it was used as a covariate in the analysis of performance (Jarrett, Glaze, Schurig, Arthur, 2010).

To determine the extent to which the study conditions influenced the outcome variables of interest, a multivariate analysis of variance (MANOVA) was conducted. The multivariate test of differences between groups using the Wilk’s Lambda criteria was not statistically significant, $F(5, 117) = 2.20$, $p > .05$. For the descriptive statistics and intercorrelations of study variables see Table 5 and Table 6, respectively.

Performance

Although the results of the MANOVA were not significant, a set of ANOVAs were conducted were conducted on each of the outcomes of interest. *H1(a)* posited that the AAR conditions (co-located and distributed) would demonstrate higher mean performance scores than the non-AAR conditions (co-located and distributed). Using a 2×3 mixed analysis of covariance (ANCOVA) with team sex composition as the covariate, the between-subject main effect was statistically significant, $F(2, 120) = 3.55$, $p < .05$, $\eta^2 = .05$ (see Table 7) indicating that the AAR teams performed better than the non-AAR teams. In addition, the within-subjects analysis indicated that teams improved across sessions, $F(3, 241) = 53.76$, $p < .05$, $\eta^2 = .30$, and there was a statistically significant training condition × session interaction, $F(3, 241) = 4.61$, $p < .05$, $\eta^2 = .03$ (see Figure 8), demonstrating that the level of performance improvement depended on the training condition (i.e., AAR vs non-AAR). Thus, *H1(a)* was supported. Additionally, analyses were conducted to determine if there were differences between the AAR and non-AAR conditions in the co-located setting. The results, $F(2, 60) = 7.01$, $p < .05$, $\eta^2 = .11$, indicated that the co-located AAR teams had higher performance score than the non-AAR teams. However, there was no significant difference between the distributed non-AAR teams and the distributed AAR teams, $F(2, 57) = 0.46$; $p > .05$, $\eta^2 = .01$.

Table 5

Means, Standard Deviations, and Intercorrelations of Study Variables

Variable	*N*	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.
1. Sex composition	120	0.53	0.26	--							
2. Team-efficacy	120	3.54	0.52	.28*	--						
3. Declarative knowledge	120	20.60	2.42	.22*	.14	--					
4. Team voice	120	4.12	0.34	.01	.49*	.14	--				
5. Team cohesion	120	4.00	0.39	.03	.57*	.11	.84*	--			
6. Reactions	79	3.52	0.42	-.08	.27*	.10	.37*	.36*	--		
7. Session 1	120	27.93	6.33	.13	-.01	.17	.09	.07	-.17	--	
8. Session 2	120	35.92	9.11	.42*	.20*	.22*	.13	.09	.07	.24*	--
9. Session 3	120	39.33	10.56	.34*	.52*	.18*	.34*	.33*	.18	.09*	.32*

Note. N is at the team level. All scales are on a 5-point Likert scale, except declarative knowledge which is scored out of points, and performance which can range from -50-100. Sex Composition indicates the number of males on a 4–person team, such that all-male = 1. * $p < .05$ (one-tailed).

Table 6

Mean and Standard Deviation of Study Variables by Training Condition

	Training Condition											
	Co-located						Distributed					
	Non-AAR (*n* = 23 teams)		Subjective AAR (*n* = 20 teams)		Objective AAR (*n* = 20 teams)		Non-AAR (*n* = 20 teams)		Subjective AAR (*n* = 20 teams)		Objective AAR (*n* = 20 teams)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Team-efficacy	3.10	0.53	3.65	0.34	3.61	0.61	3.52	0.46	3.79	0.44	3.67	0.41
Declarative knowledge	20.07	2.46	19.86	2.85	21.17	1.78	20.71	2.41	21.01	2.46	20.86	2.46
Team voice	3.80	0.30	4.24	0.28	4.23	0.25	4.04	0.43	4.30	0.26	4.08	0.25
Team cohesion	3.63	0.37	4.12	0.32	4.08	0.28	3.93	0.44	4.20	0.34	4.06	0.32
Reactions	--	--	3.67	0.27	3.72	0.45	--	--	3.47	0.38	3.41	0.40
Session 1	28.32	5.09	28.19	6.92	28.38	4.37	29.19	6.57	25.69	6.79	28.00	7.91
Session 2	31.74	9.96	35.81	9.05	36.56	9.61	36.25	8.23	35.88	9.06	38.06	8.27
Session 3	33.53	9.93	41.38	9.36	41.25	13.78	39.06	9.34	40.56	10.47	41.06	8.65

Note. The non-AAR conditions had no data for reactions. All scales are on a 5-point Likert scale, except declarative knowledge which is scored out of 30 points, and performance which can range from -50-100.

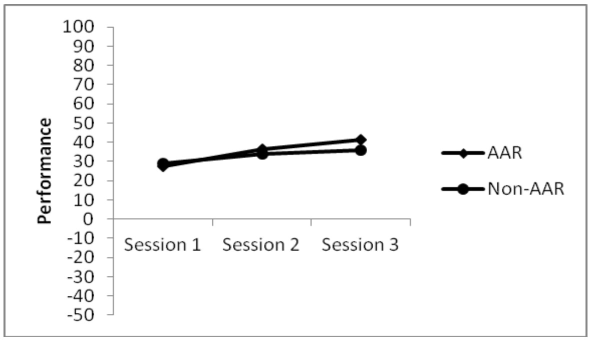

Figure 8. AAR vs non-AAR × performance session interaction.

H1(b) posited that co-located AAR teams (i.e., subjective and objective) would have higher performance scores than distributed AAR teams (i.e., subjective and objective). Using a 2×3 mixed ANCOVA with team sex composition as the covariate, the between-subject main effect was not statistically significant, $F(2, 77) = 0.11$, $p > .05$, $\eta^2 = 0.00$. Thus, *H2(b)* was not supported. The within-subjects effect was statistically significant, demonstrating improvement across the performance sessions, $F(3, 155) = 57.62$, $p < .05$, $\eta^2 = 0.42$. The training condition × session interaction was not statistically significant, $F(3, 155) = 0.42$, $p > .05$, $\eta^2 = 0.01$, indicating that performance improvement did not depend on training condition.

Table 7
Pairwise Performance Differences by Training Condition

Comparison	Session 1	Session 2	Session 3
	d	*d*	*d*
AAR vs Non-AAR	-0.20	0.29	0.48*
Co-located	-0.01	0.46*	0.72*
Distributed	-0.35	0.09	0.19
Co-lo. AAR vs Dist. AAR	0.24	-0.09	0.05
Obj. AAR vs Sub. AAR	0.21	0.14	0.02
Co-located	0.03	0.08	-0.01
Distributed	0.35	0.25	0.05

Note. Co-lo = co-located; Dist. = distributed; Obj. = objective; Subj. = subjective. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher performance scores than those in the second condition. Session 1 = mean of performance on Missions 1 and 2.Session 2 = mean of performance on Missions 3 and 4. Session 3 = mean of performance on missions 5 and 6. Total = mean of Sessions 1-3. * $p < .05$ (one-tailed).

Additionally, *H1(c)* postulated that the objective AARs teams (i.e., co-located and distributed) would have higher performance scores than subjective AAR teams (i.e., co-located and distributed). The results of the 2×3 mixed ANCOVA indicate that there was not a significant mean difference between the objective and subjective AAR teams, $F(2, 77) = 0.60$, $p > .05$, $\eta^2 = 0.01$. Hence, *H1(c)* was not supported. The within-subjects effects was statistically significant, $F(3, 155) = 57.43$, $p < .05$, $\eta^2 = 0.44$, indicating there was improvement across sessions. However, the training condition × session interaction was not statistically significant, $F(3, 155) = 0.15$, $p > .05$, $\eta^2 = 0.00$, and thus the performance improvement did not depend on the training condition. Subsequent analyses were conducted to investigate the extent to which there were performance differences for the type of AAR in either the co-located or distributed training setting. The results indicate that there were no mean differences between the co-located subjective and objective AAR teams, $F(2, 37) = 0.02$, $p > .05$, $\eta^2 = 0.00$, or the distributed subjective and objective AAR teams, $F(2, 37) = 0.86$, $p > .05$, $\eta^2 = 0.00$.

Team-Efficacy

As predicted by *H2(a),* performance was positively related to team-efficacy, $r = .52$, $p < .05$. *H2(b)* posited that teams in the co-located and distributed AAR conditions would report higher levels of team-efficacy than teams in the co-located and distributed non-AAR conditions. Using a 1-way ANOVA there was a statistically significant mean difference between the two training conditions, $F(1, 121) = 17.03$, $p < .05$, $\eta^2 = 0.12$ (see Table 8), such that teams in the AAR conditions had higher perceptions of team-efficacy than teams in the non-AAR conditions. Thus, *H2(b)* was supported. Specifically, co-located AAR teams reported higher levels of team-efficacy than non-AAR teams, $F(1, 61) = 15.03$, $p < .05$, $\eta^2 = 0.21$, but the distributed AAR teams did not report statistically significantly higher levels of team-efficacy, $F(1, 61) = 2.90$, $p > .05$, $\eta^2 = 0.05$.

Table 8
Pairwise Team-Efficacy Differences by Training Condition

Comparison	Team-efficacy *d*
AAR vs Non-AAR	0.74*
Co-located	1.04*
Distributed	0.48
Co-lo. AAR vs Dist. AAR	-0.22
Obj. AAR vs Sub. AAR	-0.18
Co-located	-0.08
Distributed	-0.28

Note. Co-lo = co-located; Dist. = distributed; Obj. = objective; Subj. = subjective. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher levels of team-efficacy than the second condition. * $p < .05$ (one-tailed).

H2(c) postulated that the co-located AAR teams (i.e., subjective and objective) would report higher levels of team-efficacy than the distributed AAR teams. The results of a 1-way ANOVA indicated the mean difference between the two training conditions was not statistically significant, $F(1, 78) = 0.93$, $p > .05$, $\eta^2 = 0.01$. Hence, *H2(c)* was not supported. Finally, *H2(d)* posited that the objective AAR teams (i.e., co-located and distributed) would report higher levels of team-efficacy than subjective AAR teams. The results of the 1-way ANOVA indicated that the mean difference between the objective and subjective AAR conditions was not statistically significant, $F(1, 78) = 0.60$, $p > .05$, $\eta^2 = 0.01$. Thus, *H2(d)* was not supported.

Declarative Knowledge

H3(a) hypothesized that declarative knowledge would be positively related to performance. The results demonstrated support for *H3(a)*, $r = .18$, $p < .05$. *H3(b)* posited that the AAR training teams (i.e., co-located and distributed, objective and subjective) would demonstrate higher declarative knowledge scores than the non-AAR teams. The results of the 1-way ANOVA indicated that there was not a statistically significant mean difference on declarative knowledge scores between the training conditions, $F(1, 121) = 0.61$, $p > .05$, $\eta^2 = 0.00$ (see Table 9). Thus, *H3(b)* was not supported.

Table 9
Pairwise Declarative Knowledge Differences by Training Condition

Comparison	Declarative Knowledge
	d
AAR vs Non-AAR	0.15
Co-located	0.18
Distributed	0.10
Co-lo. AAR vs Dist. AAR	0.18
Obj. AAR vs Sub. AAR	0.23
Co-located	0.55
Distributed	-0.07

Note. Co-lo = co-located; Dist. = distributed; Obj. = objective; Subj. = subjective. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher declarative knowledge scores than the second condition.
* $p < .05$ (one-tailed).

In addition, *H3(c)* posited that the co-located AAR teams would demonstrate higher declarative knowledge scores than the distributed AAR teams. The results of the 1-way ANOVA indicated that there was not a statistically significant mean difference between the co-located and distributed AAR teams, $F(1, 78) = 0.60$, $p > .05$, $\eta^2 = 0.01$.

Hence, *H3(c)* was not supported. Finally, *H3(d)* postulated that the objective AAR teams (i.e., co-located and distributed) would have higher declarative knowledge scores than the subjective AAR teams. Similar to the other results for declarative knowledge, there was not a statistically significant mean difference in declarative knowledge scores between the objective and subjective AAR training conditions, $F(1, 78) = 1.13$, $p > .05$, $\eta^2 = 0.01$. Thus, *H3(d)* was not supported.

Team Voice

In support of *H4(a),* there was a positive relationship between perceptions of team voice and performance, $r = .34$, $p < .05$. *H4(b)* posited that the distributed AAR teams (i.e., objective and subjective) would report higher levels of team voice than teams in the co-located AAR teams. Using a 1-way ANOVA there was not a statistically significant mean difference for perceptions of team voice between the co-located and distributed AAR teams, $F(1, 78) = 0.63$, $p > .05$, $\eta^2 = 0.01$ (see Table 10). Hence, *H4(b)* was not supported. Finally, *H4(c)* hypothesized the null, such that the perceptions of team voice for objective AAR teams were not expected to be different from those for the subjective AAR teams (i.e., co-located and distributed). The results support *H4(c)* in that there was not a statistically significant difference in levels of team voice between the objective and subjective AAR teams, $F(1, 78) = 3.56$, $p > .05$, $\eta^2 = 0.00$.

Table 10
Pairwise Team Voice Differences by Training Condition

Comparison	Team Voice
	d
Co-lo. AAR vs Dist. AAR	-0.19
Obj. AAR vs Sub. AAR	-0.42
Co-located	-0.03
Distributed	-0.57

Note. Co-lo = co-located and Dist. = distributed. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher levels of team voice than the second condition. * $p < .05$ (one-tailed).

Team Cohesion

H5(a) posited that team cohesion would be positively related to performance. The results demonstrate support for *H5(a)*, $r = .33$, $p < .05$. *H5(b)* also stated that the AAR teams (i.e., co-located and distributed, objective and subjective) would report higher perceptions of team cohesion than the non-AAR teams. The results of the 1-way ANOVA indicated that there was a statistically significant mean difference for perceptions of team cohesion across the training conditions, $F(1, 121) = 25.57$, $p < .05$, $\eta^2 = 0.17$ (see Table 11), such that the AAR teams reported higher perceptions of team cohesion than non-AAR teams. Thus, there was support for *H5(b)*. The AAR teams demonstrated higher perceptions of team cohesion than the non-AAR teams in both the co-located, $F(1, 61) = 30.18$, $p < .05$, $\eta^2 = 0.33$, and distributed, $F(1, 58) = 3.75$, $p < .05$, $\eta^2 = 0.06$, conditions.

Table 11

Pairwise Team Cohesion Differences by Training Condition

Comparison	Team Cohesion
	d
AAR vs Non-AAR	0.92*
Co-located	1.40*
Distributed	0.62*
Co-lo. AAR vs Dist. AAR	-0.09
Obj. AAR vs Sub. AAR	-0.32
Co-located	-0.13
Distributed	-0.42

Note. Co-lo = co-located; Dist. = distributed; Obj. = objective; Subj. = subjective. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher perceptions of team cohesion than the second condition. * $p < .05$ (one-tailed).

In addition, *H5(c)* posited that the co-located AAR teams would demonstrate higher perceptions of team cohesion than the distributed AAR teams. The results of the 1-way ANOVA indicated that there was not a statistically significant difference between the co-located and distributed AAR teams, $F(1, 78) = 0.21$, $p > .05$, $\eta^2 = 0.00$. Hence, *H5(c)* was not supported. Finally, *H5(d)* postulated that objective AAR teams (i.e., co-located and distributed) would not have different perceptions of team cohesion as the subjective AAR teams. The results indicate that there was not a statistically significant mean difference in perceptions of team cohesion scores between the objective and subjective AAR training conditions, $F(1, 78) = 1.83$, $p > .05$, $\eta^2 = 0.00$. Thus, *H5(d)* was not supported.

Reactions to the Training

Although there were no specific hypotheses presented for the teams’ reactions to the training, the analytical steps that were used for the other dependent variables were repeated for this outcome variable as well. There was not a statistically significant relationship between reactions to the training and performance, $r = .18$, $p > .05$. Using a 1-way ANOVA there was a statistically significant mean difference between the co-located and distributed AAR teams, $F(1, 77) = 8.88$, $p < .05$, $\eta^2 = 0.10$ (see Table 12), such that the co-located AAR teams reported more positive reactions to the AAR training. However, a 1-way ANOVA comparing the objective and subjective AAR teams indicated there were no differences in team-level reactions to the training, $F(1, 77) = 0.02$, $p > .05$, $\eta^2 = 0.00$.

Table 12
Pairwise Team-Level Reaction Differences by Training Condition

Comparison	Reactions
	d
Co-lo. AAR vs Dist. AAR	0.64*
Obj. AAR vs Sub. AAR	0.01
Co-located	0.11
Distributed	-0.18

Note. Co-lo = co-located and Dist. = distributed. *d*s were computed by subtracting the second condition from the first such that a positive *d* indicates the teams in the first condition had higher reactions to the training than the second condition. * $p < .05$ (one-tailed).

CHAPTER IV

DISCUSSION

Although limited, the extant literature provides favorable support for the ability of AARs to improve cognitive, skill, and attitude-based outcomes. The objective of this study was to investigate the comparative effectiveness of different types of AAR training (i.e., objective and subjective) in co-located and distributed conditions on a variety of training effectiveness outcomes. Specifically, comparisons were made across a variety of team-related outcome variables (i.e., performance, team-efficacy, declarative knowledge, team voice, team cohesion, and reactions to the training).

The results of this study indicated that the teams in the AAR conditions (i.e., co-located and distributed, objective and subjective) had higher mean performance scores and demonstrated a faster rate of performance improvement than teams in the non-AAR training conditions. Moreover, AAR teams reported higher levels of team-efficacy and team cohesion than the non-AAR teams. Finally, the co-located AAR teams reported more favorable reactions to the training than the distributed AAR teams. However, the objective AARs were no more effective than the subjective AARs, regardless of whether they were co-located or distributed, in terms of performance, declarative knowledge, team-efficacy, team voice, team cohesion, or reactions to the training.

Study Objectives

The present study advances the team training literature via three primary objectives. The first objective was to investigate the comparative effectiveness of co-located and distributed AAR training on a variety of outcome variables. To the best of this author's knowledge, this study provided the first comparative evaluation of AAR training in co-located and distributed training environments. Although previous studies have investigated the effectiveness of distributed AAR training protocols (i.e., Kring 2004; Oden, 2009), there has yet to be a study that investigates the distributed subjective and objective AAR's efficacy compared to co-located subjective and objective AAR training protocols.

The second objective was to provide a more comprehensive examination of the effectiveness of AAR training across performance, knowledge, emergent states, and attitude-based outcomes. That is, there have been other studies that have *separately* investigated performance (e.g., Ellis & Davidi, 2005; Kring, 2004; Oden, 2009; Villado & Arthur, 2012), declarative knowledge, team-efficacy (Villado & Arthur, 2012), and team cohesion (Kring, 2004). Using a variety of outcome variables to assess training interventions is consistent with the extant training literature which indicates that a training manipulation may have differential effectiveness based on the outcome of interest (Arthur, Bennett, Stanush, & McNelly, 1998; Schmidt & Bjork, 1992). For instance, Schmidt and Bjork (1992) demonstrated that training interventions may be capable of enhancing participant effectiveness during the acquisition phase, but not for more distal outcomes such as long-term retention. Thus, this study sought to examine a variety of outcome variables under the same conditions to determine the extent to which AAR training can influence these outcomes of interest. In addition, the inclusion of team voice and reactions to the training in the present study provides additional outcome variables that have not yet been examined in the AAR literature.

The final objective was to provide additional evidence pertaining to the effectiveness of subjective and objective AAR training. In practice, training designers continue to increase the objectivity of the AAR training via the increased use of technology to capture and record performance episodes for review. However, the empirical research that has investigated the benefits of objective AAR has been somewhat inconsistent with the theoretical models posed by researchers. For example, Villado and Arthur (2012) found that subjective and objective AAR reviews were equally capable of influencing performance, declarative, knowledge and team-efficacy. In addition, Savoldelli et al (2006) obtained similar results as Villado and Arthur.

There remain many instances where recording performance may be cumbersome or monetarily unfeasible and thus, understanding the value added, via objective reviews, allows one to better estimate the utility of objective reviews. It is not unreasonable to posit that objective AARs provide little to no incremental validity over and above subjective AARs, which would suggest that in scenarios where objective AARs are not feasible, subjective AARs may be an equally effective alternative. Thus, the present study sought to provide additional evidence that speaks to the effectiveness of objective AAR reviews across different training settings (i.e., co-located and distributed).

AAR vs Non-AAR Training

In congruence with research demonstrating that AAR teams (either objective or subjective) reported higher perceptions of team-efficacy than the non-AAR teams (Villado & Arthur, 2012; Villado et al., 2011), the present study also found that AAR teams reported higher levels team-efficacy and team cohesion than the non-AAR teams.

Thus, whereas AAR training would commonly be considered a taskwork intervention that is focused on improving task-related knowledge and/or performance (Kozlowski & Bell, 2003), the present study illustrates that the AAR can have unintended favorable effects on teamwork skills as well. This finding is of interest to team training researchers and practitioners given that some researchers would argue that both taskwork and teamwork skills are necessary for a team to demonstrate superior performance (Salas et al., 1998). Hence, although AAR training is focused on improving taskwork skills, by providing an opportunity for individuals to openly discuss team-related issues, teams may also develop and improve their teamwork skills (e.g., communication skills, team cohesion).

As previously noted, the primary focus of the AAR in the present study was to increase task-related performance. Consistent with the extant literature that has investigated the effectiveness of AAR training (e.g., Ellis & Davidi, 2005; Ellis et al., 2010; Ellis et al., 2009; Kring, 2004; Schurig et al., 2011; Villado & Arthur, 2012), the present study found that teams that received AAR training not only demonstrated higher performance scores over baseline, but they also had higher performance scores and a faster rate of performance improvement than teams who did not receive AAR training (i.e., non-AAR conditions). These findings would seem to provide support for the continued use of the AAR as a training method. However, it is noteworthy that the AAR in the present study did not influence team declarative knowledge acquisition. Thus, the AAR's use in situations where knowledge acquisition is the proximal outcome of interest may necessitate adjustments to the process in order to provide maximal effectiveness.

One potential caveat to the finding of mean differences between the AAR and non-AAR conditions is that when one examines the mean differences within the two geographic settings (i.e., co-located and distributed) there is a significant mean difference between the co-located AAR and non-AAR conditions, but this difference does not persist in the distributed AAR vs non-AAR comparison. That is, it would seem that the mean performance differences in the co-located conditions are driving the results of the omnibus test. The mean differences between the co-located and distributed AAR conditions were small (*d*s ranged from -.09 - .24) and thus the lack of a significant main effect for the distributed AAR vs non-AAR comparison may be a result of the small sample size. In summary, the results indicate that the AAR is potentially an effective training intervention in co-located environments, but the results for the distributed environment are less conclusive.

The lack of a significant finding between the distributed AAR and non-AAR teams could be a result of the reduced trust between the geographically dispersed team members (Kring, 2004; Rosen et al., 2006). The nature of the geographic distribution, wherein one set of tank operators (Tank 1) were geographically dispersed from the second set of tank operators (Tank 2), could have led to faultlines—dividing lines based on the alignment of one or more group member attributes (Lau & Murnighan, 1998)—that may have adversely impacted the AAR process. These faultlines may have led to more difficulty gaining consensus on the effective and ineffective actions as well as process loss when developing a strategy. There may have been more internal discussions between tank operators and less across the two tanks leading to more difficulty performing the mission effectively. Given the high levels of interdependence necessary to perform the task, any impediments to communication may have made mission success more challenging.

Co-located vs Distributed AAR Training Environments

The inconsistent findings between the co-located and distributed AAR teams, concomitant with the increased prevalence of geographically distributed training, demonstrates the importance of empirically investigating the AAR's effectiveness in distributed environments. A recent empirical review of this literature concluded that co-located teams are superior to distributed teams in terms of both performance and affective outcomes (Baltes et al., 2002) and consequently, concluded that technology-mediated teams are rarely more effective than co-located teams. For example, Kring (2004) found that the teams in the co-located AAR condition had higher levels of performance than the teams in the distributed AAR condition. However, in contrast to previous research using action teams that demonstrated the superiority of co-located teams over distributed teams, the present study found no significant differences between the teams in co-located and distributed AAR conditions.

One potential explanation for the lack of a significant difference between co-located and distributed AAR teams is the instructional design properties of the AAR. That is, meta-analytic evidence demonstrates that the instructional design of the training intervention moderates the relationship between co-located and distributed performance

(Sitzmann, Kraiger, Stewart, & Wisher, 2006). For example, Sitzmann et al. (2006) found that when the training method incorporated learner control and feedback, there was no performance difference between co-located and distributed training. However, when examining training methods that did not incorporate these instructional design elements, there were significant performance differences between the co-located and distributed training conditions (Sitzmann et al., 2006). This evidence suggests that the level of structure in the training intervention (e.g., AAR) may directly influence the extent to which there are performance differences between co-located and distributed training. Thus, the structure of the AAR (Villado, 2008) may be a critical design feature in mitigating the performance differences between co-located and distributed AAR teams.

It was hypothesized that given the reduced amount of communication cues and greater levels of anonymity, teams participating in distributed AARs would demonstrate reduced knowledge acquisition, reduced team cohesiveness, and lower perceptions of team-efficacy. However, with the exception of team-level reactions, there were no significant mean differences on any of the outcome variables of interest. That is, teams that participated in a distributed AAR had similar levels of team-efficacy, team cohesion, and declarative knowledge as the co-located AAR teams. However, this finding is not completely at odds with previous literature as there is some evidence in the extant AAR literature which indicates that AAR training is equally effective regardless of the type of AAR (e.g., self-led vs instructor-led AARs, objective vs subjective AARs, level of accuracy; Ellis et al., 2009; Villado & Arthur, 2012; Villado et al., 2011).

It is important to note that distributed AAR teams demonstrated significantly lower reactions to the training, compared to the co-located AAR teams. This represents the first study, to this author's knowledge, that has investigated mean differences in team-level trainee reactions. The use of team-level reactions provides insight into how team-level reactions are related to other team-level attitudinal and behavioral outcomes. Team-level reactions to training could influence their motivation to perform the training. In addition, training reactions are also positively related to training transfer (Alliger et al., 1997) and thus, teams that vary greatly on their training reactions may have greater difficulty transferring the skills learned in training to new environments. Evaluating team-level reactions allows researchers to use similar techniques found in the team composition literature and investigate team-level relationships with greater precision.

To the extent that team reactions are important to the training developers, whether it is to increase motivation, engagement, or training transfer, developers should be cognizant of the potentially lower levels of reactions to distributed training that may be present. However, if team reactions are not a primary focus, such as situations in which there are other forms of motivation to perform the training successfully (e.g., compliance or compensation), it may provide practitioners with a more conducive environment for the use of AAR interventions in distributed environments.

Objective vs Subjective AAR Training

Practitioners have invested large amounts of resources to integrate technology into the review process in order to increase the objectivity of the AAR review with the supposition that it will provide greater benefit to the AAR participants. As such, the importance of objective AAR review systems and the effectiveness of intrinsic vs extrinsic feedback have been identified as design characteristics that warrant further research (Villado, 2008). Yet, the empirical research that has investigated this topic has been inconclusive. Villado and Arthur (2012) posited that the subjective AAR conditions would have larger amounts of errors of omission and commission and would subsequently lead to less benefits garnered from subjective AAR teams. However, they found that the type of AAR review (objective or subjective AAR) did not influence team performance. Consistent with the results of Villado and Arthur (2012), the results of the present study indicate that there were no significant mean differences between objective and subjective AARs on any of the outcome variables of interest.

Thus, it would seem that inaccuracies in the performance review portion of the AAR does not result in team process loss. That is, there may be some metacognitive benefits to simply participating in the AAR regardless of the extent to which the performance review is accurate (Ellis et al., 2009) or even related to the team's *own* performance (Ellis et al., 2010). Metacognition consists of two components, knowledge acquisition—accumulated knowledge about cognitive process—and regulation of knowledge which is the use of regulatory strategies to facilitate cognitive performance (Baker & Brown, 1984; Flavell, 1979). Thus, AAR training may represent a type of knowledge regulation and subsequently influence learning and performance by facilitating the cognitive process, regardless of the type of AAR. These findings suggest that although practitioners consider objective AAR review systems to be more favorable in improving outcomes, the empirical evidence suggests that there may be relatively little utility to the use of objective AAR reviews over subjective AARs.

In addition to the potential metacognitive effects of the AAR, it seems that the type of error in recall may be an important moderating factor when establishing potential differences between objective and subjective AAR reviews. Specifically, when individuals overestimated their performance (as opposed to underestimating), the error was less detrimental to performance improvement than when individuals underestimated the performance in the previous mission (Ellis et al., 2009). Thus, the proposition that inaccuracies in the subjective review process would result in lower performance may have been amiss. Future research should catalogue the amount and type of inaccuracies in the subjective review to provide a more accurate assessment of the potential influence of inaccuracies on future performance.

Another potential explanation for the absence of differences between the objective and subjective AAR teams on the outcome variables, is that given the massed nature of the protocols (five hours, similar to Villado & Arthur, 2012) there is the common criticism that studies using teams performing complex tasks do not have adequate time to develop task proficiency (Ackerman & Cianciolo, 2000) or learn to work as a team. That is, novices may be overwhelmed by the amount of information presented in the objective AAR and conversely as team members become experts they may be more capable of processing objective information specific to their previous performance. Thus, in this situation it may be possible that the subjective and objective AARs are functioning quite similarly and trainees are not capable of garnering the benefits of the objective AAR. Determining if teams at varying levels of the skill acquisition phases (Ackerman, 1987) are capable of utilizing the AAR information differently represents an additional design characteristic (e.g., expert vs novice) that needs to be examined.

Related to the massed nature of the protocol, the spacing of the AAR or proximity of the AAR to the performance episode represents a potentially important factor when examining the efficacy objective AARs. The present study had the teams perform the AAR directly after the performance episode, so teams were more likely to have the details of the previous performance episode accessible. However, in situations where there may be long gaps between the performance episode and the AAR, the use of objective AARs may be more beneficial because they do not rely on the teams' ability to recall critical incidents from the performance episode. The longer the time interval between the performance episode and the review, the lower the likelihood that teams will be able to accurately assess their performance without the help of memory aids (e.g., recordings or diaries). Thus, additional research is needed to determine how the proximity of the AAR to the performance episode affects the extent to which objective AARs can provide benefits to trainees, and indeed even be more differentiated from subjective AARs in terms of their effectiveness.

Practical and Scientific Implications

The findings of the present study highlight several practical and scientific implications that should be considered when using or investigating AAR training. Primarily, it would seem that regardless of the geographic dispersion or type of AAR review (e.g., objective vs subjective), AAR training remains an effective intervention in terms of increasing performance and certain attitudinal-based outcomes. Thus, their continued use should provide benefits to teams and individuals across a variety of outcome variables (e.g., performance, team-efficacy, team cohesion). However, the present study's findings coupled with those of Villado and Arthur (2012) demonstrate that AARs like that used in the present study may not be effective for improving declarative knowledge acquisition. This is not to say that AAR training cannot provide benefits to knowledge acquisition, but if this is the goal of researchers and practitioners, then the training design should ensure that the AAR elicits declarative knowledge information during the review process. However, it is important to note that these conclusions are being drawn from a limited number of studies. Thus, additional research is needed to determine the extent to which AAR training can be designed to improve knowledge acquisition in individuals and teams.

Although the results comparing the distributed and co-located AARs were somewhat inconsistent, the findings of this study seem to suggest that the use of distributed AARs (as opposed to co-located) do not engender the proposed process loss that was hypothesized. Thus, the results of this study indicate that the use of distributed training to reduce administrative costs, without any process loss, may be a viable option for organizations that have geographically dispersed individuals and teams. However, the one caveat is that teams in the distributed AAR conditions had lower reactions to the AAR training than teams in the co-located AAR conditions. Thus, to the extent that training reactions are a primary concern, it is important to note the difference between the co-located and distributed training reactions.

The current findings provide an important initial assessment of the efficacy of distributed AAR training protocols. However, the results here are contrary to those reported by Kring (2004) and given the inconsistent evidence, the use of AARs in co-located and distributed teams should continue to be a design characteristic of interest to practitioners and researchers (Villado, 2008). Given the importance of team training to organizations, additional research beyond what has been reviewed and investigated here is needed to more conclusively establish the effect of distributed training on performance and other outcomes of interest (Salas et al., 2002).

Similarly, practitioners should evaluate the extent to which the use of technology to capture and record performance for a more objective performance review (e.g., subjective vs objective AAR reviews), provides the intended benefit to the trainees. The empirical research has consistently demonstrated that the use of objective or personal review systems provides little to no benefit to the trainees. Concomitant with the limited benefits is the increased cost of developing and implementing the technology. Further research is needed to determine whether objective AARs are beneficial to the organization. For instance, additional research is needed to determine whether experts are better capable of utilizing all of the information presented to them in an objective review or if the proximity of the AAR to the performance episode is an important factor to consider when designing the AAR review system.

Limitations and Future Directions

As with any study, there are some limitations associated with the present study. One such limitation was the use of undergraduate college students assigned to ad hoc teams. Sears (1986) proposes that the limited subject pool used in most psychological research (i.e., mostly white, middle class undergraduates) restricts our ability to develop a generalizable set of theoretical propositions. This serves as a potential external validity threat because many training interventions involving teams typically involve intact teams. However, Mook (1983) points out that many research questions are specifically interested in non-generalizable effects as a way to test what can potentially happen, or to test inferred processes observed in field settings. Thus, this study represents an initial assessment of distributed AAR training and should be validated using samples that are more consistent with the population of interest. Further research is needed to determine the extent to which these findings generalize to other populations and specifically to the organizations that implement these sorts of training interventions.

A second potential limitation, related to the preceding, was the use of a massed training protocol (five hours; Arthur et al., 2010). For example, Arthur et al. (2010) found that the length of the interstudy time interval (in a multi-stage study) was positively related to immediate posttraining performance and long-term skill retention. The use of a massed protocol may have masked any effects between the conditions. Consequently, a training design that has a longer inter-session interval may be more effective in detecting differences between the training conditions. The potential impact of skill acquisition phases on generalizability has also been noted by Yeo and Neal

(2006); they note that results from short-term skill acquisition studies may not generalize to long-term skill acquisition. This presents an important area of future research, either using organizational samples or long-term lab-based designs, to examine whether the findings of this study (and others like it) generalize to actions teams that have repeated training trials and are together for longer periods of time.

Another potential limitation was the use of a self-led AAR as opposed to instructor-led AARs, which are probably more common in field and applied settings. However, self-led AARs were used to maintain a systematic procedure and reduce the extent to which experimenter effects and bias could influence the AAR process. It is also noteworthy that Villado et al. (2011) found that there were no significant mean differences between self-led AAR training and instructor-led AAR training. Thus, this issue may not be a particularly major limitation as would otherwise seem to be the case. Furthermore, it is not unreasonable to posit that some organizations use self-led AARs as a training intervention (e.g., deployed combat teams). Nevertheless, additional research is needed to determine the extent to which the findings obtained in the present study would persist if the AAR was instructor-led instead of self-led.

There is increasing research interest in the AAR, largely in part to keep pace with its use in practice. However, previous research only provides initial insights into the effectiveness and boundary conditions of AAR training and additional research is needed to understand the generalizability and applicability of this training intervention in different settings and using different design features. Thus, as previously noted, there remain several AAR characteristics that warrant new and additional research. Expanding on and adding to Villado (2008)’s list, these characteristics include: (1) the number of trainees (team size), (2) the training of complex versus simple tasks, (3) the provision of intrinsic versus extrinsic feedback (Alexander et al., 1962), (4) the training of individuals versus teams, (5) whether the performance episode reviewed was a successful or failed performance episode (Ellis & Davidi, 2005, 2006), (6) whether training is co-located versus distributed (the present study; Kring, 2004; Oden, 2009), (7) the frequency of the AAR, (8) the spacing of the AAR, (9) the degree of structure imposed on the AAR (Ellis et al., 2010; Prince et al., 2005), (10) whether the AAR is self-led or instructor- or facilitator-led (Villado et al., 2011), (11) the objectivity versus subjectivity of the AAR (Savoldelli et al., 2006; Villado & Arthur, 2012), and (12) the time interval between the performance episode and the AAR. Thus, whereas certain characteristics have received *some* research attention, there remain several characteristics that have yet to be researched and it is clear that without additional research to understand how these design features influence the relationship between the AAR and the outcomes of interest, questions regarding the AARs effectiveness will remain.

CHAPTER V

CONCLUSION

The team-training literature provides favorable support for the AAR's ability to improve cognitive, skill, and attitudinal outcomes in co-located and distributed environments. However, the comparative effectiveness of co-located and distributed AARs is relatively unknown. Thus, the objective of the present study was to investigate the comparative effectiveness of co-located and distributed AARs across a multitude of outcome variables (i.e., performance, declarative knowledge, team-efficacy, team voice, team cohesion, and team-level reactions).

The study's hypotheses (see Table 1) were tested using 4-person teams performing an interdependent complex team task. The results indicated that teams in the AAR conditions had significantly higher performance scores than the teams in the non-AAR conditions. In addition, the AAR teams reported higher perceptions of team-efficacy and higher levels of team cohesion than the non-AAR teams. With the exception of team-level reactions, there were no significant differences between the distributed AAR and co-located AAR teams on any of the outcomes of interest. Similarly, there were no significant differences across any of the outcome variables between the objective and subjective AAR conditions, indicating that the type of AAR did not impact the results of the training.

The findings of the present study highlight several potential practical and scientific implications that could be considered when using or investigating AAR training. Primarily, regardless of the geographic dispersion or type of AAR (e.g., objective vs subjective), it would seem that AAR training remains an effective intervention at increasing performance and attitudinal-based outcomes (e.g., team-efficacy, team cohesion). In addition, the results suggest that the use of distributed AARs (as opposed to co-located) does not engender the proposed process loss that one would have expected. Thus, the use of distributed AARs to reduce administrative costs, without a commensurate loss in effectiveness, may be a viable option for organizations with geographically dispersed teams and individuals. Finally, practitioners should evaluate the extent to which the use of technology to monitor performance is beneficial to the trainees. The lab-based, empirical research has consistently demonstrated that the use of objective or personal review systems provides little to no benefit to the trainees. Future research is needed to determine the generalizability of these findings to other tasks, domains, team types, and levels of expertise.

REFERENCES

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences in skill acquisition. *Journal of Experimental Psychology: Applied, 6,* 259-290.

Alexander, L. T., Kepner, C. H., & Tregoe, B. B. (1962). The effectiveness of knowledge of results in a system-training program. *Journal of Applied Psychology, 46,* 202-211.

Allen, T. J. (1977). *Managing the flow of technology.* Cambridge, MA: MIT Press.

Alliger, G. M., Tannenbaum, S. I., Bennett, W. Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of relations among training criteria. *Personnel Psychology, 50,* 341-358.

Ancona, D. G., & Caldwell, D. F. (1992). Demography and design: Predictors of new product team performance. *Organization Science, 3,* 321-341.

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89,* 369-406.

Arthur, W. Jr., Bell, S. T., & Edwards, B. D. (2007). A longitudinal examination of the comparative criterion-related validity of additive and referent-shift consensus operationalizations of team-efficacy. *Organizational Research Methods, 10,* 35-58.

Arthur, W. Jr., Bennett, W., Jr., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 83,* 234-245.

Arthur, W. Jr., Bennett, W. Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative analysis. *Human Performance, 11,* 57–101.

Arthur, W. Jr., Edwards, B. D., Bell, S. T., & Bennett, W. Jr. (2002). *Identifying factors that influence the cognitive readiness and performance of teams: The development and evaluation of a high fidelity synthetic team task performance environment (STTPE) lab*. Grant/contract No. F41624-97-C-5000, L-3 Communications Corporation, Link Simulation and Training Division, Mesa, AZ.

Arthur, W. Jr., Edwards, B. D., Bell, S. T., Villado, A. J., & Bennett, W. Jr. (2005). Team task analysis: Identifying tasks and jobs that are team based. *Human Factors, 47*, 654-659.

Arthur, W. Jr., Glaze, R. M., Bhupatkar, A., Villado, A. J., Bennett, W. Jr., & Rowe, L. J. (2012). Team task analysis: Differentiating between tasks and using team relatedness and team workflow as metrics of team task interdependence. *Human Factors, 54*, 277-295.

Baird, L., Holland, P., & Deacon, S. (1999). Learning from action: Imbedding more learning into the performance fast enough to make a difference. *Organizational Dynamics, 27,* 19-32.

Baker, D. P., & Salas, E. (1996). Analyzing team performance: In the eye of the beholder? *Military Psychology, 8,* 235-245.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp.353-394). New York: Longman.

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., & LaGanke, J. S. (2002). Computer-mediated communication and group decision making: A meta-analysis. *Organizational Behavior and Human Decision Process, 87,* 156-179.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Banks, A. P., & Millward, L. J. (2007). Differentiating knowledge in teams: The effect of shared declarative and procedural knowledge on team performance. *Group Dynamics: Theory, Research, and Practice, 11,* 95-106.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*, 377-391.

Barry, B., & Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: The role of personality. *Journal of Applied Psychology, 82,* 62-78.

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). Cohesion and performance in groups: A meta-analytic clarification of construct relations. *Journal of Applied Psychology, 88,* 989-1004.

Bordia, P. (1997). Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. *Journal of Business Communication, 34,* 99-120.

Bouas, K. S., & Arrow, H. (1996). The development of group identity in computer and face-to-face groups with membership change. *Computer Supported Cooperative Work, 4,* 153-178.

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing communication sequences for team training needs assessment. *Human Factors, 40,* 672-679.

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology, 54,* 271-296.

Brown, K. G., & Sitzmann, T. (2011). Training and employee development for improved performance. In Zedeck (Ed.), *APA handbook of industrial and organizational psychology, vol 2: Selecting and developing members for the organization* (pp. 469-503). Washington, DC, US: American Psychological Association.

Brown, S. L., & Eisenhardt, K. M. (1995). Product development: Past research, present findings, and future directions. *Academy of Management Review, 20,* 343-378.

Cannon-Bowers, J. A., & Salas, E. (1997). Teamwork competencies: The interaction of team member knowledge, skills, and attitudes. In H. F. O'Neil (Eds.), *Workforce readiness: Competencies and assessment* (pp. 45-62). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cannon-Bowers, J. A., Salas, E., Tannenbaum, S. I., & Mathieu, J. E. (1995). Toward theoretically based principles of training effectiveness: A model and initial empirical investigation. *Military Psychology, 7,* 141-164.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333-380). San Francisco: Jossey-Bass.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.). *Perspectives on socially shared cognition* (pp. 127-151). Washington D.C.: American Psychological Association.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior, 6*, 191-223.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 554-571.

Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection and management performance: Implications for information systems. *MIS Quarterly, 11*, 355-366.

DeNisi, A. S., Robbins, T., & Cafferty, T. P. (1989). Organization of information used for performance appraisals: Role of diary-keeping. *Journal of Applied Psychology, 74*, 124-129.

DeNisi, A. S., & Sonesh, S. (2011). The appraisal and management of performance at work. In Zedeck (Ed.), *APA handbook of industrial and organizational psychology, vol 2: Selecting and developing members for the organization* (pp. 255-279). Washington, DC, US: American Psychological Association.

Dennis, A. R., & Valacich, J. S. (1999). Rethinking media richness: Towards a theory of media synchronicity. *System Sciences, 32,* 1-10.

Detert, J. R., & Burris, E. R. (2007). Leadership behavior and employee voice: Is the door really open? *Academy of Management Journal, 50*, 869-884.

Devine, D. (1999). Effects of cognitive ability, task knowledge, information sharing, and conflict on group decision-making effectiveness. *Small Group Research, 30*, 608-634.

Devine, D. J., Clayton, L. D., Philips, J. L., Dunford, B. B., & Melner, S. B. (1999). Teams in organizations prevalence, characteristics, and effectiveness. *Small Groups Research, 30*, 678-711.

Devine, D. J., & Phillips, J. L. (2001). Do smarter teams do better: A meta-analysis of cognitive ability and team performance. *Small Group Research*, *32*, 507-532.

Driskell, J. E., Radtke, P. H., & Salas, E. (2003). Virtual teams: Effects of technological mediation on team performance. *Group Dynamics: Theory, Research, and Practice, 7,* 297-323.

Dwyer, D. J., Oser, R. L., Salas, E., & Fowlkes, J. E. (2000). Performance measurement in distributed environments: Initial results and implications for training. *Military Psychology, 11*, 189-215.

Edwards, B. D., Day, E., Arthur, W. Jr., & Bell, S. T. (2006). Relationships among team ability composition, team mental models, and team performance. *Journal of Applied Psychology, 91*, 727-736.

Ellis, A. P. J., Bell, B. S., Ployhart, R. E., Hollenbeck, J. R., & Ilgen, D. R. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. *Personnel Psychology, 58,* 641-672.

Ellis, S., & Davidi, I. (2005). After-event reviews: Drawing lessons from successful and failed experience. *Journal of Applied Psychology, 90,* 857-871.

Ellis, S., Ganzach, Y., Castle, E., & Sekely, G. (2010). The effect of filmed versus personal after-event reviews on task performance: The mediating and moderating role of self-efficacy. *Journal of Applied Psychology, 95,* 122-131.

Ellis, S., Mendel, R., & Aloni-Zohar, M. (2009). The effect of accuracy of performance evaluation on learning from experience: The moderating role of after-event reviews. *Journal of Applied Social Psychology, 39,* 541-563.

Ellis, S., Mendel, R., & Nir, M. (2006). Learning from successful and failed experiences: The moderating role of kind of after-event review. *Journal of Applied Psychology, 91,* 669-680.

eSim Games. (2007). *Steel Beasts Pro PE ver. 2.370* [Computer software]. Mountain View, CA: Author.

Eys, M. A., & Carron, A. V. (2001). Role ambiguity, task cohesion, and task self-efficacy. *Small Groups Research, 32,* 356-373.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Feldman, D. C. (1989). Socialization, resocialization, and training: The role of attitude accessibility. In I. L. Goldstein & Associates (Eds.) *Training and development in* *organizations* (pp.376-414). San Francisco, CA: Jossey Boss.

Festinger, L. (1950). Informal social communication. *Psychological Review*, *57*, 271-284.

Fiore, S. M., Salas, E., Cuevas, H. M., & Bowers, C. A. (2003). Distributed coordination space: Toward a theory of distributed team process and performance. *Theoretical Issues in Ergonomic Science, 4,* 340-364.

Fjermestad, J., & Hiltz, S. R. (1999). An assessment of group support system experimental research: Methodology and results. *Journal of Management Information Systems, 15,* 7-149.

Flavell, J. E. (1979). Metacognition and cognitive monitoring: A new area of cognitive development inquiry. *American Psychologist, 34,* 906-911.

Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & O ser, R. (1994). Improving the measurement of team performance: The TARGET methodology. *Military Psychology, 6,* 47-61.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Toronto, Ontario, Canada: Wadsworth.

Goodwin, G. F. (1999). *Predicting performance in U. S. Army Special Forces teams: Staffing and composition issues*. Unpublished dissertation. The Pennsylvania State University.

Gully, S. M., Devine, D. J., & Whtiney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of levels of analysis and task interdependence. *Small Groups Research, 26,* 497-520.

Gully, S. M., Incalcaterra, K. A., Joshi, A., & Beaubien, J. M. (2002). A meta-analysis of team-efficacy, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology, 87,* 819-832.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law, 2,* 447-472.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69,* 85-98.

Jarrett, S. M., Glaze, R. M., Arthur, W. Jr., Schurig, I., Villado, A. J., & Bennett, W. Jr. (2010). *The role of sex composition in team training performance.* Paper presented at the 25th annual meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Jessup, L. M., Connolly, T., & Galegher, J. (1990). The effects of anonymity on GDSS group process with an idea-generating task. *Management Information Systems Quarterly, 14*, 313-321.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training and Development, 13,* 3-9.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (2nd ed., pp.

301-319. New York: McGraw-Hill.

Kirkpatrick, D. L. (1996). Invited reaction: Reaction to the Holton article. *Human Resource Development Quarterly, 7,* 23-25.

Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Vol. 12. Industrial and organizational psychology* (pp. 333-375). London: Wiley.

Kozlowski, S. W. J., & Bell, B. S. (2007). A theory-based approach for designing distributed learning systems. In S. M. Fiore & E. Salas (Eds.). *Towards a science of distributed learning,* (pp.15-39). Washington, DC: American Psychological Association.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311-328.

Kring, J. P. (2004). *Communication modality and after action review performance in a distributed immersive virtual environment* (Unpublished doctoral dissertation). University of Central Florida, Orlando, Florida.

Lau, D., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review, 23,* 325-340.

Lepine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with Big

Five personality characteristics and cognitive ability. *Journal of Applied Psychology, 86,* 326-336.

Lindell, M. K., Brandt, C. J., & Whitney, D. J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement, 23*, 127-135.

Lindsley, D. H., Brass, D. J., & Thomas J. B. (1995). Efficacy-performance spirals: A multi-level perspective. *Academy of Management Review, 20,* 645-678.

Littlepage, G., Robinson, W., & Reddington, K. (1997). Effects of task experience and group experience on group performance, member ability, and recognition of expertise. *Organizational Behavior and Human Decision processes, 69,* 133-147.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57,* 705-717.

Lynch, L. M., & Black, S. E. (1998). Beyond the incidence of employer-provided training. *Industrial & Labor Relations Review, 52,* 64-81.

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefing and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85,* 971-986.

Markus, M. L. (1994). Electronic mail as a medium of managerial. *Organization Science, 5,* 502-527.

Mathieu, J. E., Heffner, T. S., Goodwin, G. E., Cannon-Bowers, J. A., & Salas, E. (2005). Scaling the quality of teammates' mental models: equifinality and normative comparisons. *Journal of Organizational Behavior, 26*, 37-56.

Meliza, L. L., Bessemer, D. W., & Hiller, J. H. (1994). Providing unit training feedback in the distributed interactive simulation environment. In R. F. Holz, J. H. Hiller, & H. H. McFann (Eds.) *Determinants of effective unit performance: Research on measuring and managing unit training readiness* (pp. 257-280). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Mento, A. J., Steel, R. P., & Karen, R. J. (1987). A meta-analytic study of goal setting on task performance: 1966-1984. *Organizational Behavior and Human Decision Processes, 39*, 52-83.

Mesmer-Magnus, J. R., Dechurch, L. A., Jimenez-Rodriguez, M., Wildman, J., & Shuflller, M. (2011). A meta-analytic investigation of virtuality and information sharing in teams. *Organizational Behavior and Human Decision Processes, 115,* 214-225.

Middendorf, C. H., & Macan, T. H. (2002). Note-taking in the employment interview: Effects on recall and judgments. *Journal of Applied Psychology, 87*, 293-303.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379-387.

Murphy, K. R. (1991). Criterion issues in performance appraisal research: Behavioral accuracy versus classification accuracy. *Organizational Behavior and Human Decision Processes, 50,* 45-50.

Murphy, K. R., & Balzer, W. K. (1986). Systematic distortions in memory-based ratings and performance evaluations: Consequences for rating accuracy. *Journal of Applied Psychology, 71,* 39-44.

Noe, A. R. (2010). *Employee training and development* (5th ed.). New York: Mcgraw Hill.

Oden, K. B. (2009). *Distributed team training: Effective team feedback.* (Unpublished doctoral dissertation). University of Central Florida, Orlando, Florida.

Olivera, F., & Straus, S. G. (2004). Group-to-individual transfer of learning cognitive and social factors. *Small Groups Research, 35,* 440-465.

Oser, R. L., Gualtieri, J. W., Cannon-Bowers, J. A., & Salas, E. (1999). Training team problem solving skills: an event-based approach. *Computers in Human Behavior, 15,* 441-462.

Porter, C. O. L. H. (2005). Goal orientation: Effects on backing up behavior, performance, efficacy, and commitment in teams. *Journal of Applied Psychology, 90,* 811-818.

Prince, C., Salas, E., Brannick, M., & Orasnau, J. (2005). Beyond facilitation: An improved CRM debrief for safety training. *Human Factors and Aerospace Safety, 5,* 1-22.

Rapp, T. L., & Mathieu, J. E. (2007). Evaluating an individually self-administered generic teamwork skills training program across time and levels. *Small Groups Research, 38,* 532-555.

Rice, R. E., & Shook, D. (1990). Relationships of job categories and organizational levels to use of communication channels, including electronic mail: A meta-analysis and extension. *Journal of Management Studies, 27,* 195-229.

Rico, R., & Cohen, S. G. (2005). Effects of task interdependence and type of communication on performance in virtual teams. *Journal of Managerial Psychology, 20,* 261-274.

Roche, L., Kurt, C., & Ahrens, S. (2010). An avatar approach to distributed team training for mission operations. *Acta Astronautica, 67,* 1164-1169.

Ron, N., Lipshitz, R., & Popper, M. (2006). How organizations learn: Post-flight reviews in an F-16 fighter squadron. *Organizational Studies*, *27,* 1069-1089.

Rosen, B., Furst, S., & Blackburn, R. (2006). Training for virtual teams: An investigation of current practices and future needs. *Human Resource Management, 45,* 229-247.

Ryan, A. M., Daum, D., Bauman, T., Grisez, M., Mattimore, K., Nalodka, & McCormick, S. (1995). Direct, indirect, and controlled observation and rating accuracy. *Journal of Applied Psychology, 80,* 664-670.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology, 8,* 197-208.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors, 50,* 903-933.

Salas, E., Oser, R. L., Cannon-Bowers, J. A., & Daskarolis-Kring, E. (2002). Team training in virtual environments: An event based approach. In M. K. Stanney (Ed.) *Handbook of virtual environments: Design, implementation, and applications. Human factors and ergonomics* (pp. 873-892). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Sanchez, J. I., & De La Torre, P. (1996). A second look at the relationship between rating and behavioral accuracy in performance appraisal. *Journal of Applied Psychology, 81,* 3-10.

Savoldelli, G. L., Naik, V. N., Park, J., Joo, H. S., Chow, R., & Hamstra, S. J. (2006). Value of debriefing during simulated crisis management: Oral versus video-assisted oral feedback. *Anesthesiology, 105,* 279-285.

Schurig, I., Jarrett, S., Arthur, W. Jr., Glaze, R. M., & Schurig, M. (2011). *The effectiveness of after-action reviews as a training method: A meta-analysis.* Unpublished manuscript, College Station, TX: Texas A&M University.

Schmidt, R. A., & Björk, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207-217.

Sears, J. T. (1986). *Conducting qualitative research in higher education*. Presented at the International Conference of First Year Experience in Newcastle upon Tyne, 1986.

Sia, C. L., Tan, B. C. Y., & Wei, K. K. (2002). Group polarization and computer-mediated communication: Effects of communication cues, social presence, and anonymity. *Information Systems Research, 13,* 70-90.

Sitzmann T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008) A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology, 93,* 280-293.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness and classroom instruction: A meta-analysis. *Personnel Psychology, 59,* 623-664.

Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., & Salas, E. (2008). Guided team self-correction: Impacts on team mental models, processes, and effectiveness. *Small Group Research, 39,* 303-327.

Spears, R., & Lea, M. (1994). Panacea or panopticon: The hidden power in computer-mediated communication. *Communication Research, 21,* 427-459.

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management, 37,* 503-530.

Straus, S. G., & Olivera, F. (2000). Knowledge acquisition in virtual teams. In M. A. Neale & E. A. Mannix & T. L. Griffith (Eds.). *Research in managing groups and teams: Vol. 3. Technology* (pp. 257-282). Stamford, CT: JAI.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors, 41,* 61-71.

Sturman, M. C., Cheramie, R. A., & Cashen, L. H. (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology, 90,* 269-283.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

Townsend, A. M., Demarie, S. M., & Hendrickson, A. R. (1998). Virtual teams: Technology and the workplace of the future. *Academy of Management Executive, 12*, 17-29.

Triana, M. C., Kirkman, B. L., & Wagstaff, M. F. (2011). Does the order of face-to-face and computer-mediated communication matter in diverse project teams? An investigation of communication order effects on minority inclusion and participation. *Journal of Business and Psychology, 27,* 57-70.

Tubbs, M. E. (1986). Goal setting: A meta-analytic examination of the empirical evidence. *Journal of Applied Psychology, 71,* 474-483.

United States Army (1993). *A leader's guide to after-action reviews* (Training Circular 25-20), Washington, DC.

Van Dyne, L., & Lepine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal, 41*, 108-119.

Villado, A. J. (2008). *The after-action-review training approach: An integrative framework and empirical investigation.* Unpublished dissertation, College

Station, TX: Texas A&M University.

Villado, A. J., & Arthur, W. Jr. (2012). *An empirical investigation of the effectiveness of subjective and objective after-action reviews in a complex task training environment.* Manuscript submitted for publication.

Villado, A. J., Arthur, W. Jr., & Bennett, W. Jr. (2009). *The after-action review approach: An empirical test.* Paper presented at the 24th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Villado, A. J., Cosio, C. P., Pawlik, C. E., Rice, A. Y., Narain, P. O., & Baron, A. C. (2011). *The expert-led after-action review training approach: An empirical test.* Paper presented at the 26th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Villado, A. J., Day, E. A., Arthur, W. Jr., Boatman, P. R., Kowollik, V., Bhupatkar, A., & Bennett, W. Jr. (in press). Use of, reaction to, and efficacy of observation rehearsal training: Enhancing skill retention on a complex command-and-control simulation task. In W. Arthur, Jr., E. A. Day, W. Bennett, Jr., & A Portrey (Eds.), *Individual and team skill decay: The science, and implications for practice.* New York: Taylor & Francis/Psychology Press.

Warkentin, M., Sayeed, L., & Hightower, R. (1999). Virtual teams versus face-to-face teams: An exploratory study of web-based conference system. In K. E. Kendall (Ed.), *Emerging information technologies: Improving decisions, cooperation, and infrastructure* (pp. 241-262). Thousand Oaks, CA: Sage Publications

Webb, N. M. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz & N. Miller (Eds.). *Interaction in cooperative groups: The theoretical anatomy of group learning.* (pp.102-119). New York: Cambridge University Press.

Whiting, S. W., Podsakoff, P. M., & Peirce, J. R. (2008). Effects of task performance, helping, voice, and organizational loyalty on performance appraisal ratings. *Journal of Applied Psychology, 93,* 125-139.

Williams, K. J., DeNisi, A. S., Meglino, B. M., & Cafferty, T. P. (1986). Initial decisions and subsequent performance ratings. *Journal of Applied Psychology, 71*, 189-195.

Woehr, D. J., & Feldman, J. (1993). Processing objective and question order effects on the causal relation between memory and judgment in performance appraisal: The tip of the iceberg. *Journal of Applied Psychology, 78,* 232-241.

Yeo, G. B., & Neal, A. (2006). An examination of the dynamic relationship between self-efficacy and performance across levels of analysis and levels of specificity. *Journal of Applied Psychology, 91,* 1088-1101.

Zakay, D., Ellis, S., & Shevaliski, M. (2004). Outcomes' values and early warning indications as determinants of willingness to learn from experience. *Experimental Psychology, 51*, 150-157.

APPENDIX A

EXAMPLE ITEMS FROM THE DECLARATIVE KNOWLEDGE MEASURE

STEEL BEASTS DECLARATIVE KNOWLEDGE TEST

Research ID: ____

For each question, please select the alternative that you think is the **BEST** answer. Mark you answer by filling in the bubble next to your selection. Please treat this test as you would a classroom exam. You are allowed to look at your keyboard and joystick, but you are not allowed to use any other materials nor discuss the answers with others—this is a "closed book" test.

Please complete all items on this test. There is no penalty for guessing so it is in your best interest to guess if you do not know the answer to a question.

1. The JOYSTICK hat button is used by the tank commander to:

   ○A. designate targets for the gunner.
   ○B. fire the .50 caliber coaxial gun.
   ○C. set the view relative to the turret.

2. Which of the following JOYSTICK buttons should be pressed by the gunner to toggle between 3× and 10× magnification?

   ○A. 2
   ○B. 4
   ○C. 6

3. When the map DISPLAY is set to LOS, the white area on the map indicates points at which:

   ○A. a majority of your tank is fully exposed.
   ○B. only the turret of your tank is exposed.
   ○C. your tank is fully exposed to thermal imaging.

4. Which of the following tanks is the enemy using?

   ○A. Leopard 1A4
   ○B. M1A1 Abrams
   ○C. T-72 Ural

5. The tank commander must press the ___________ KEYBOARD key to request that the loader begin loading HEAT ammunition.

   ○A. DELETE
   ○B. HOME
   ○C. INSERT

6. Which of the following functions is disrupted when the stabilization is damaged?

   ○A. Aiming the main gun
   ○B. Driving the tank
   ○C. Firing the main gun

7. Routes may be created by the tank commander on the:

   ○A. eye view screen.
   ○B. map screen.
   ○C. route toolbar.

APPENDIX B

TEAM-EFFICACY MEASURE

TEAM-EFFICACY

Research ID: ____

Please read each of the statements listed below and mark the response that best indicates the extent to which you agree with each statement

①	②	③	④	⑤
Strongly Disagree	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**

Rate each of the following statements to indicate the extent to which they are descriptive of **YOUR** opinions of your **PLATOON**.

1.	I feel confident in **my platoon's** ability to perform well on Steel Beasts.	① ② ③ ④ ⑤
2.	I think **my platoon** can meet the challenges of Steel Beasts.	① ② ③ ④ ⑤
3.	I know **my platoon** can achieve good scores on Steel Beasts.	① ② ③ ④ ⑤
4.	I know **my platoon** can master Steel Beasts.	① ② ③ ④ ⑤
5.	I do **NOT** think Steel Beasts is something that **my platoon** will become good at.	① ② ③ ④ ⑤
6.	I am confident that **my platoon** has what it takes to perform well on Steel Beasts.	① ② ③ ④ ⑤

APPENDIX C

TEAM VOICE AND TEAM COHESION MEASURES

TEAM VOICE AND TEAM COHESION

Research ID: ____

Please read each of the statements listed below and mark the response that best indicates the extent to which you agree with each statement

① Strongly Disagree	② Disagree	③ Neither Agree nor Disagree	④ Agree	⑤ Strongly Agree

1.	My platoon members and I had a chance to express our opinions	① ② ③ ④ ⑤
2.	My platoon members and I enjoyed interacting with each other	① ② ③ ④ ⑤
3.	My platoon members and I trusted each other	① ② ③ ④ ⑤
4.	My platoon members and I felt free to make negative comments	① ② ③ ④ ⑤
5.	My platoon members and I would like to have future teams include similar members	① ② ③ ④ ⑤
6.	My platoon members and I worked well together	① ② ③ ④ ⑤
7.	My platoon members and I felt free to make positive comments	① ② ③ ④ ⑤
8.	My platoon members and I listened to each others' inputs	① ② ③ ④ ⑤
9.	My platoon members and I liked each other	① ② ③ ④ ⑤
10.	My platoon members and I shared a feeling of cohesion within our platoon	① ② ③ ④ ⑤

APPENDIX D

DEMOGRAPHICS

DEMOGRAPHICS

Please read each of the statements listed below and respond in a manner that best describes you.

Research ID:____

Age: __________

Sex: ☐ Male ☐ Female

Race: ____

Major: ____

Occupation: ____

Highest education earned (check one):

Doctoral Degree	☐ Completed	☐ In-Progress			
Master's Degree	☐ Completed	☐ In-Progress			
Bachelor's Degree	☐ Completed	☐ Senior	☐ Junior	☐ Sophomore	☐ Freshman
Associates Degree	☐ Completed	☐ In-Progress			
Technical/Vocational	☐ Completed	☐ In-Progress			
High School	☐ Completed	☐ 12th grade	☐ 11th grade	☐ 10th grade	☐ 9th grade

Generally, what is your playing ability regarding video/computer games? (check one)

☐ Novice ☐ Average ☐ Expert

What is your playing ability on *Steel Beasts*™? (check one)

☐ Never Played ☐ Average ☐ Expert

APPENDIX E

TEAM-LEVEL REACTIONS MEASURE

TEAM-LEVEL REACTIONS

Research ID: ____

Please read each of the statements listed below and mark the response that best indicates the extent to which you agree with each statement

①	②	③	④	⑤
Strongly Disagree	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**

1.	Participating in the after-action-reviews makes me feel uncomfortable.	① ② ③ ④ ⑤
2.	Participating in the after-action-reviews is frustrating.	① ② ③ ④ ⑤
3.	I enjoy participating in the after-action-reviews.	① ② ③ ④ ⑤
4.	Participating in the after-action-reviews is confusing.	① ② ③ ④ ⑤
5.	I am bored participating in the after-action-reviews.	① ② ③ ④ ⑤
6.	Participating in the after-action-reviews helps me learn to play *Steel Beasts*.	① ② ③ ④ ⑤
7.	Participating in the after-action-reviews motivates me to try harder.	① ② ③ ④ ⑤
8.	Participating in the after-action-reviews helps me improve my *Steel Beasts* performance.	① ② ③ ④ ⑤
9.	Participating in the after-action-reviews shows me strategies for playing *Steel Beasts* that I was unaware of.	① ② ③ ④ ⑤
10.	Participating in the after-action-reviews makes me confident about my ability to play *Steel Beasts*.	① ② ③ ④ ⑤
11.	I would have learned to play *Steel Beasts* much better **without** the after-action-reviews.	① ② ③ ④ ⑤
12.	If I had the opportunity to participate in this study again, I would prefer to play *Steel Beasts* **without** the after-action-reviews.	① ② ③ ④ ⑤

VITA

Steven Jarrett

Office: Texas A&M University
Department of Psychology
College Station, Texas 77843-4235

Phone: (979)-412-1348
Email: sjarrett1@gmail.com

Education

Ph.D August 2012 — Industrial/Organizational Psychology
Texas A&M University, College Station, Texas.

M.S. December 2010 — Industrial/Organizational Psychology
Texas A&M University, College Station, Texas.

B.S. May 2006 — Psychology, Organizational and Leadership Strategies, and Portuguese dual minor
Purdue University, West Lafayette, Indiana

Refereed Journal Articles

Beus, J. M., **Jarrett, S. M.,** Payne, S. C., & Bergman, M. E. Perceptual equivalence of psychological climates across organizational faultlines: When agreement indices do not agree. *Journal of Occupation and Organizational Psychology.* [in press]

Manuscripts Under Review

Glaze, R. M., **Jarrett, S. M.,** Schurig, I., Arthur, W. Jr., & Taylor, J. E. Comparative evaluation of three situational judgment test response formats in terms of construct-related validity, subgroup differences, and socially desirable responding. *Personnel Psychology*.

Jarrett, S. M., Glaze, R., Schurig, I., Arthur, W. Jr. The importance of team sex composition in team training research employing complex psychomotor tasks. *Human Performance.*

Schurig, I., **Jarrett, S. M.,** Glaze, R. M., & Arthur, W. Jr. A meta-analysis of the effectiveness of after-action reviews as a training method. *Human Factors* [revise-and-resubmit]